Pharmacological treatment for psychotic depression

Wijkstra, Jaap; Lijmer, Jeroen; Burger, Huibert; Cipriani, Andrea; Geddes, John; Nolen, Willem A

Published in:
Cochrane database of systematic reviews (Online)

DOI:
10.1002/14651858.CD004044.pub4

Cochrane Library

Cochrane Database of Systematic Reviews

Pharmacological treatment for psychotic depression (Review)

Wijkstra J, Lijmer J, Burger H, Cipriani A, Geddes J, Nolen WA

Wijkstra J, Lijmer J, Burger H, Cipriani A, Geddes J, Nolen WA.
Pharmacological treatment for psychotic depression.
Cochrane Database of Systematic Reviews 2015, Issue 7. Art. No.: CD004044.
DOI: 10.1002/14651858.CD004044.pub4.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Pharmacological treatment for psychotic depression

Jaap Wijkstra[1], Jeroen Lijmer[2], Huibert Burger[3,4], Andrea Cipriani[5,6], John Geddes[7], Willem A Nolen[4]

[1]Department of Psychiatry, UMCU, Utrecht, Netherlands. [2]Department of Psychiatry, Onze Lieve Vrouwe Gasthuis, Amsterdam, Netherlands. [3]Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands. [4]Department of Psychiatry, University Hospital Groningen, Groningen, Netherlands. [5]Department of Public Health and Community Medicine, University of Verona, Verona, Italy. [6]Department of Psychiatry, Univeristy of Oxford, Oxford, UK. [7]Department of Psychiatry, University of Oxford/Warneford Hospital, Oxford, UK

Contact address: Willem A Nolen, Department of Psychiatry, University Hospital Groningen, Groningen, Netherlands. w.a.nolen@psy.umcg.nl.

Editorial group: Cochrane Common Mental Disorders Group.
Publication status and date: New search for studies and content updated (conclusions changed), comment added to review, published in Issue 7, 2015.

Citation: Wijkstra J, Lijmer J, Burger H, Cipriani A, Geddes J, Nolen WA. Pharmacological treatment for psychotic depression. *Cochrane Database of Systematic Reviews* 2015, Issue 7. Art. No.: CD004044. DOI: 10.1002/14651858.CD004044.pub4.

ABSTRACT

Background

Evidence is limited regarding the most effective pharmacological treatment for psychotic depression: combination of an antidepressant plus an antipsychotic, monotherapy with an antidepressant or monotherapy with an antipsychotic. This is an update of a review first published in 2005 and last updated in 2009.

Objectives

1. To compare the clinical efficacy of pharmacological treatments for patients with an acute psychotic depression: antidepressant monotherapy, antipsychotic monotherapy and the combination of an antidepressant plus an antipsychotic, compared with each other and/or with placebo.

2. To assess whether differences in response to treatment in the current episode are related to non-response to prior treatment.

Search methods

A search of the Cochrane Central Register of Controlled Trials and the Cochrane Depression, Anxiety and Neurosis Group Register (CCDANCTR) was carried out (to 12 April 2013). These registers include reports of randomised controlled trials from the following bibliographic databases: EMBASE (1970-), MEDLINE (1950-) and PsycINFO (1960-). Reference lists of all studies and related reviews were screened and key authors contacted.

Selection criteria

All randomised controlled trials (RCTs) that included participants with acute major depression with psychotic features, as well as RCTs consisting of participants with acute major depression with or without psychotic features, that reported separately on the subgroup of participants with psychotic features.

Data collection and analysis

Two review authors independently extracted data and assessed risk of bias in the included studies, according to the criteria of the *Cochrane Handbook for Systematic Reviews of Interventions*. Data were entered into RevMan 5.1. We used intention-to-treat data. For dichotomous efficacy outcomes, the risk ratio (RR) with 95% confidence intervals (CIs) was calculated. For continuously distributed outcomes, it was not possible to extract data from the RCTs. Regarding the primary outcome of harm, only overall dropout rates were available for all studies.

Main results

The search identified 3659 abstracts, but only 12 RCTs with a total of 929 participants could be included in the review. Because of clinical heterogeneity, few meta-analyses were possible. The main outcome was reduction of severity (response) of depression, not of psychosis.

We found no evidence for the efficacy of monotherapy with an antidepressant or an antipsychotic.

However, evidence suggests that the combination of an antidepressant plus an antipsychotic is more effective than antidepressant monotherapy (three RCTs; RR 1.49, 95% CI 1.12 to 1.98, P = 0.006), more effective than antipsychotic monotherapy (four RCTs; RR 1.83, 95% CI 1.40 to 2.38, P = 0.00001) and more effective than placebo (two identical RCTs; RR 1.86, 95% CI 1.23 to 2.82, P = 0.003).

Risk of bias is considerable: there were differences between studies with regard to diagnosis, uncertainties around randomisation and allocation concealment, differences in treatment interventions (pharmacological differences between the various antidepressants and antipsychotics) and different outcome criteria.

Authors' conclusions

Psychotic depression is heavily understudied, limiting confidence in the conclusions drawn. Some evidence indicates that combination therapy with an antidepressant plus an antipsychotic is more effective than either treatment alone or placebo. Evidence is limited for treatment with an antidepressant alone or with an antipsychotic alone.

PLAIN LANGUAGE SUMMARY

Pharmacological treatment for psychotic depression

Psychotic depression is a severe depression with psychotic features (i.e. delusions and/or hallucinations). Uncertainty surrounds the most effective drug treatment for psychotic depression: with an antidepressant alone, with an antipsychotic alone or with the combination of an antidepressant plus an antipsychotic.

The aim of this review is to compare the efficacy of the various forms of drug treatment that have been used to treat psychotic depression. We did this by analysing all randomised controlled trials (RCTs) that investigated drug treatments for psychotic depression. We searched for these trials in a wide-ranging way. The search identified 3659 studies, but in the end, we found only 12 RCTs that met our inclusion criteria. These trials involved a total of 929 people.

From these trials, we found evidence that the combination of an antidepressant plus an antipsychotic provides more effective treatment for psychotic depression than either treatment alone. However, our confidence in this conclusion is limited because the information came from only a small number of RCTs, which included small numbers of people. In addition, the types of people involved varied between RCTs, and the RCTs differed in design, which means that we cannot confidently generalise their findings.

BACKGROUND

Description of the condition

Psychotic depression is a severe condition that is defined as a depressive episode with psychotic features (i.e. delusions and/or hallucinations) in the context of a (unipolar) major depressive disorder. Psychotic depression is not uncommon. In the US Epidemiologic Catchment Area Study (Johnson 1991), 14% of participants who met the criteria for major depression had a history of episodes with psychotic features. In a European general population study, 18.5% of respondents with a major depressive episode had psychotic features; the prevalence of psychotic depression was 0.4% and of non-psychotic depression 2.0% (Ohayon 2002). In a US study of hospitalised participants with major depression, 25% met the criteria for psychotic depression (Coryell 1984). Compared with non-psychotic depression, psychotic depression is marked by greater severity, increased incapacity, a lower likelihood of placebo response, longer duration of episodes and recurrence of psychotic features in subsequent episodes (Coryell 1998).

Description of the intervention

Guidelines (APA 2010; NICE 2009) recommend electroconvulsive therapy (ECT) or pharmacotherapy as treatment for psychotic depression. Pharmacotherapy for psychotic depression could consist of an antipsychotic, an antidepressant or a combination of both. Most guidelines recommend treatment that combines an antidepressant with an antipsychotic (APA 2010; NICE 2009. However, discussion continues regarding whether the combination of an antipsychotic plus an antidepressant is more effective than monotherapy with an antidepressant or an antipsychotic (Dutch Guideline 2009; Mahli 2009; Parker 1992; Wijkstra 2005; Wijkstra 2007). The intervention studied in this review is pharmacological treatment for psychotic depression, especially the question of whether the combination of an antipsychotic plus an antidepressant is more effective than either treatment given as monotherapy.

How the intervention might work

All antidepressants enhance the activity of serotonin and/or noradrenaline, and some of them (also) dopamine (Sadock 2009). Most antidepressants achieve this via inhibition of reuptake of these neurotransmitters in the presynaptic neuron (tricyclic antidepressants (TCAs), selective serotonin reuptake inhibitors (SSRIs) and serotonin-noradrenaline reuptake inhibitors (SNRIs)), although some antidepressants have other working mechanisms (e.g. blockade of postsynaptic serotonin-2 receptors such as mirtazapine or inhibition of their breakdown via inhibition of the enzyme monoamine oxidase (MAOIs)). Nevertheless, their noradrenergic and serotonergic effects do not completely explain their efficacy, as these effects occur already within hours after first intake, but it takes days to weeks before antidepressants begin to exert their effects in participants with depression or anxiety (Sadock 2009).

Almost all antipsychotics (classical as well as atypical antipsychotics, with the exception of clozapine) are blockers of the postsynaptic dopamine-2 receptor, and their therapeutic efficacy is correlated with their affinity for dopamine-2 receptors in vivo. However, other effects may contribute to their efficacy, such as their affinity for presynaptic serotonin-1 receptors, postsynaptic serotonin-2 receptors and histamine receptors, as can be seen with some atypical antipsychotics (e.g. olanzapine and quetiapine). Similar to the antidepressants, these effects do not completely explain their efficacy because they also occur already within hours after first intake, but it takes days to weeks for antipsychotics to begin to work (Sadock 2009).

The traditional view is that antidepressants work against depression and antipsychotics work against psychosis. Therefore, it seems appropriate in psychotic depression to treat the psychotic symptoms with an antipsychotic and the depressive symptoms with an antidepressant. However, when psychotic depression is considered as the most severe form of depression, and when psychosis is viewed as the distal consequence of that severity (as is the case in the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR);* APA 2000), treatment with an antidepressant alone seems logical. On the other hand, treatment with an antipsychotic alone, especially one of the newer atypical antipsychotics with possibly antidepressant effects, cannot be ruled out.

Why it is important to do this review

Clinical practise is characterised by uncertainty as to whether it is most appropriate to start treatment in this patient group with antidepressant monotherapy or with the combination of an antidepressant and an antipsychotic because of the potential adverse effects of antipsychotics (especially extrapyramidal side effects, hyperprolactinaemia and the risk of metabolic syndrome, including weight gain). A previous meta-analysis did not find a statistically significant difference between TCA monotherapy and combination therapy (Parker 1992). However, the findings and conclusions of that meta-analysis were limited by the inadequate methodology of many of the included studies, which were often retrospective, uncontrolled and/or not randomised. Some international guidelines on the pharmacological treatment of psychotic depression (in the United States: Nelson 1997; in the Netherlands: Dutch Guideline 2009) and those presented in reviews (Wheeler 2000) suggest that one may consider TCA monotherapy before adding an antipsychotic. However, in contrast, the American Psychiatric Association (APA) Practice Guideline for the Treatment of Pa-

tients with Major Depressive Disorder (APA 2010) and the National Institute for Health and Clinical Excellence (NICE 2009) recommend initial combination therapy. The same recommendation is made in a review by Coryell (Coryell 1998). This variation between guidelines reflects the limited evidence on which these guidelines are based. In a review about evidence used in practice guidelines (Wijkstra 2007), we concluded that physicians (and patients) should be aware that guidelines for treatment recommendations may be less evidence-based than asserted, even when it is stated that treatment recommendations are based on the highest level of evidence.

Treatment with a classical antipsychotic alone is not recommended, primarily because of the findings of a study by Spiker (Spiker 1985), in which treatment with perphenazine alone was less effective than treatment with perphenazine plus amitriptyline. However, the atypical antipsychotics may be worth reconsidering now because of the reduced risk of extrapyramidal side effects and potential antidepressant properties of some of the atypical antipsychotics (Rothschild 2004a).

This review is an update of our Cochrane review first published in 2005 (Wijkstra 2005). Our conclusion in 2005 was as follows: "Treatment with an antipsychotic alone is not a good option. Starting with the combination of an antidepressant and an antipsychotic, as well as starting with an antidepressant alone and adding an antipsychotic if the patient does not respond, both appear to be appropriate options for patients with psychotic depression."

Since 2005, a few more studies have been conducted, leading to a different conclusion regarding treatment with an antidepressant alone or with the combination of an antidepressant and an antipsychotic.

Another important clinical issue is that differences in response to specific treatments may be explained in relation to non-response to prior treatment(s). Generalising from observations across medical disciplines, it would be expected that patients who did not respond to an adequate treatment will respond less to subsequent treatment. Some data are available on this topic with regard to pharmacological treatment of major depressive disorder (Sackeim 2001). Two studies (Prudic 1990; Prudic 1996) showed that a greater degree of treatment resistance predicts an inferior response to ECT.

OBJECTIVES

1. To compare the clinical efficacy of pharmacological treatments for patients with an acute psychotic depression: antidepressant monotherapy, antipsychotic monotherapy and the combination of an antidepressant plus an antipsychotic, compared with each other and/or with placebo.

2. To assess whether differences in response to treatment in the current episode are related to non-response to prior treatment.

METHODS

Criteria for considering studies for this review

Types of studies

We included randomised controlled trials (RCTs) of the pharmacological treatment of participants with acute psychotic depression.

As we expected to identify very few RCTs assessing the treatment of psychotic depression as the primary focus, we decided a priori to also include RCTs assessing the treatment of major depression with or without psychotic features. Effects in the subgroup of participants with psychotic features should then be reported separately, irrespective of whether the subgroup with psychotic features was stratified before randomisation.

No language restrictions were applied for included studies.

Types of participants

Participants were of any age in any setting (both inpatients and outpatients) who had a major depressive disorder and a current episode with psychotic features (delusions and/or hallucinations appearing in the context of a full major depressive episode) according to the *Diagnostic and Statistical Manual of Mental Disorders, Third Edition (DSM-III)/DSM, Third Edition, Revised (DSM-III-R)/DSM, Fourth Edition (DSM-IV)/DSM, Fourth Edition, Text Revision (DSM-IV-TR)* (APA 1980; APA 1987; APA 1994; APA 2000) or *International Classification of Diseases (ICD)* codes for the same.

We also included studies in which participants had comorbidities, as comorbidity was not a reason for exclusion.

As patients with a major depressive episode in the context of a bipolar disorder (bipolar depression) are at increased risk of switching to mania (Licht 2008), we intended to include only trials that assessed participants with unipolar depression. If trials had included participants with both unipolar and bipolar depression, we decided a priori to include the trial only when results in the unipolar group were reported separately, or when the percentage of participants with bipolar depression did not exceed 20% of the total study population.

Types of interventions

We included any pharmacological treatment of a current (i.e. acute) episode. Treatment had to be given for at least four weeks with the intention of treating the current episode.

Where possible, the following pairwise comparisons were considered.

1. Antidepressant versus placebo.
2. Antipsychotic versus placebo.
3. Antidepressant versus antidepressant.

4. Antipsychotic versus antipsychotic.
5. Antidepressant versus antipsychotic.
6. Antidepressant plus antipsychotic versus placebo.
7. Antidepressant plus antipsychotic versus placebo plus antidepressant.
8. Antidepressant plus antipsychotic versus placebo plus antipsychotic.

Types of outcome measures

Primary outcomes

1. Efficacy: clinical response of depression based on observer-rated symptom reduction: reduction of at least 50% on the Hamilton Rating Scale for Depression (HRSD, Hamilton 1960) or the Montgomery Åsberg Depression Rating Scale (MADRS, Montgomery 1979) or any other observer depression severity rating scale, or a change score on the Clinical Global Impression-Change (CGI-C, Guy 1976) of 'much improved' or 'very much improved'.
2. Harm: overall dropout rate during acute treatment as a proxy measure of overall acceptability of treatment.

Secondary outcomes

3. Remission of depression as defined in the reports and based on HRSD or MADRS or other observer depression severity rating scale.
4. Change from baseline in score on the HRSD, MADRS or any other observer depression severity rating scale or change in severity on Clinical Global Impression-Severity (CGI-S).
5. Quality of life, as defined in the reports.
6. Dropout rate due to adverse effects.

Search methods for identification of studies

CCDAN's Specialised Register (CCDANCTR)

The Cochrane Depression, Anxiety and Neurosis Group (CCDAN) maintains two clinical trials registers at its editorial base in Bristol, UK; a references register and a studies-based register. The CCDANCTR-References Register contains over 31,500 reports of trials in depression, anxiety and neurosis. Approximately 65% of these references have been tagged to individual, coded trials. The coded trials are held in the CCDANCTR-Studies Register and records are linked between the two registers through the use of unique Study ID tags. Coding of trials is based on the EU-Psi coding manual. Reports of trials for inclusion in the Group's registers are collated from routine (weekly), generic searches of MEDLINE, EMBASE and PsycINFO; quarterly searches of the Cochrane Central Register of Controlled Trials (CENTRAL) and review specific searches of additional databases. Reports of trials are also sourced from international trials registers c/o the World Health Organization's trials portal (ICTRP), drug companies, and the handsearching of key journals, conference proceedings and other (non-Cochrane) systematic reviews and meta-analyses. Details of CCDAN's generic search strategies can be found on the Group's website.

Electronic searches

The CCDANCTR-Studies Register was searched (all years to 12 April 2013) using the following terms:
Condition = (depressi* or "affective disorder*" or "affective symptoms")
AND
Condition or Comorbidity = (psychosis or psychoses or psychotic* or delusion*)
The CCDANCTR-Studies Register was searched (all years to 12 April 2013) using the following terms to identify additional untagged references:
Title/Abstract/Keywords = ((depressi* or "affective disorder*" or "affective symptoms")
AND
Free-Text=(psychosis or psychoses or psychotic* or delusion* or hallucin* or antipsychotic* or psychotropic*))
An additional search of the Cochrane Central Register of Controlled Trials (CENTRAL) was carried out (Issue 4, 2010) (Appendix 1).
A summary of searches carried out for the earlier version of this review can be found in Appendix 2.

Searching other resources

The reference lists of all studies, related reviews and relevant conference proceedings were screened and key authors contacted.

Data collection and analysis

Selection of studies

In step 1, all abstracts of identified publications were screened independently by two review authors (JL and JW), and studies were selected if they met the following criteria.
1. RCT.
2. Included participants with a major depressive disorder.
3. Investigated the efficacy of pharmacological treatment.
4. Concerned acute phase treatment (minimum of four weeks treatment), not continuation or maintenance treatment.

If any doubt or disagreement arose between the review authors, the publication was included in step 2. Full articles were obtained for the selected abstracts.
In step 2, selected full articles were screened (JL and JW) according to the following criteria.
1. Participants with psychotic depression were not excluded.

2. Results in the subgroup of psychotic depressed participants were reported separately.

If any doubt arose about the article, it was included in step 4.

In step 3, the reference lists of related reviews and of included publications, conference abstracts and personal communications were searched.

Finally, in step 4, two review authors (JL and JW) independently reviewed all identified publications according to the full inclusion criteria of the review. Any disagreement was resolved by consensus discussion with another review author (WN).

Data extraction and management

Extracted data included the following: participant characteristics (age, gender, setting: inpatients/outpatients); diagnosis (which diagnostic instrument, system of classification, number of bipolar participants); prior treatment of the current episode; intervention; length of illness; suicide attempts; treatment details (treatment period, washout period, additional medication, blood levels, doses) and outcome measures. Data were extracted independently by two review authors (JW and JL).

All data (from 2004 and recent data) were entered into RevMan 5.1 (RevMan 2011).

Assessment of risk of bias in included studies

In the original version of this review, we assessed methodological quality of included studies by criteria set out in the *Cochrane Handbook for Systematic Reviews of Interventions* (Alderson 2004); however, after publication of the revised and expanded *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), we updated our methods accordingly. Working independently, JW and JL assessed risk of bias of included studies using the tool described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). The following items were assessed.

1. Sequence generation: Was the allocation sequence adequately generated?
2. Allocation concealment: Was allocation adequately concealed?
3. Blinding of participants, personnel and outcome assessors for each main outcome or class of outcomes: Was knowledge of the allocated treatment adequately prevented during the study?
4. Incomplete outcome data for each main outcome or class of outcomes: Were incomplete outcome data adequately addressed?
5. Selective outcome reporting: Were reports of the study free of suggestion of selective outcome reporting?
6. Other sources of bias: Was the study apparently free of other problems that could put it at high risk of bias?

We included quotations from the text of included studies and comments on how we assessed the risk of bias; we judged each study to be of low, unclear or high risk of bias.

See risk of bias figures (Figure 1; Figure 2).

Figure 1. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

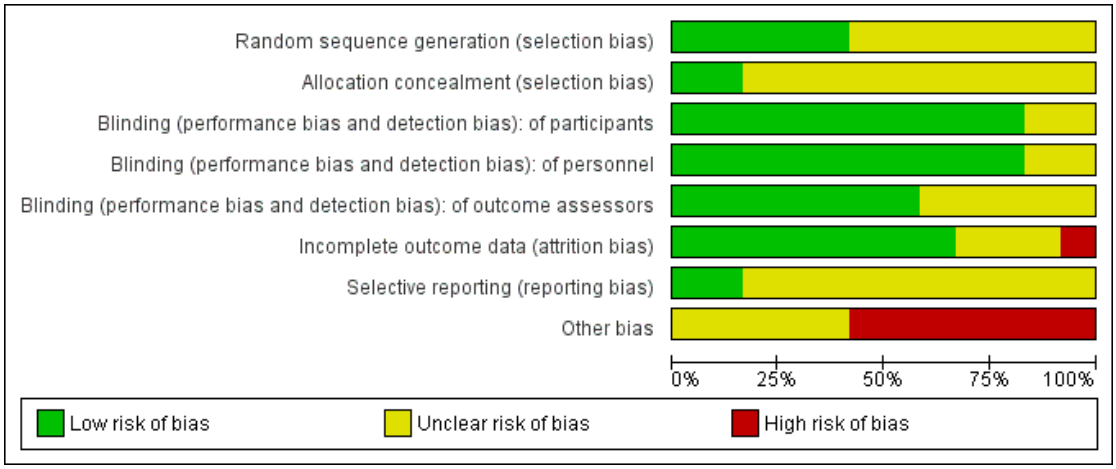
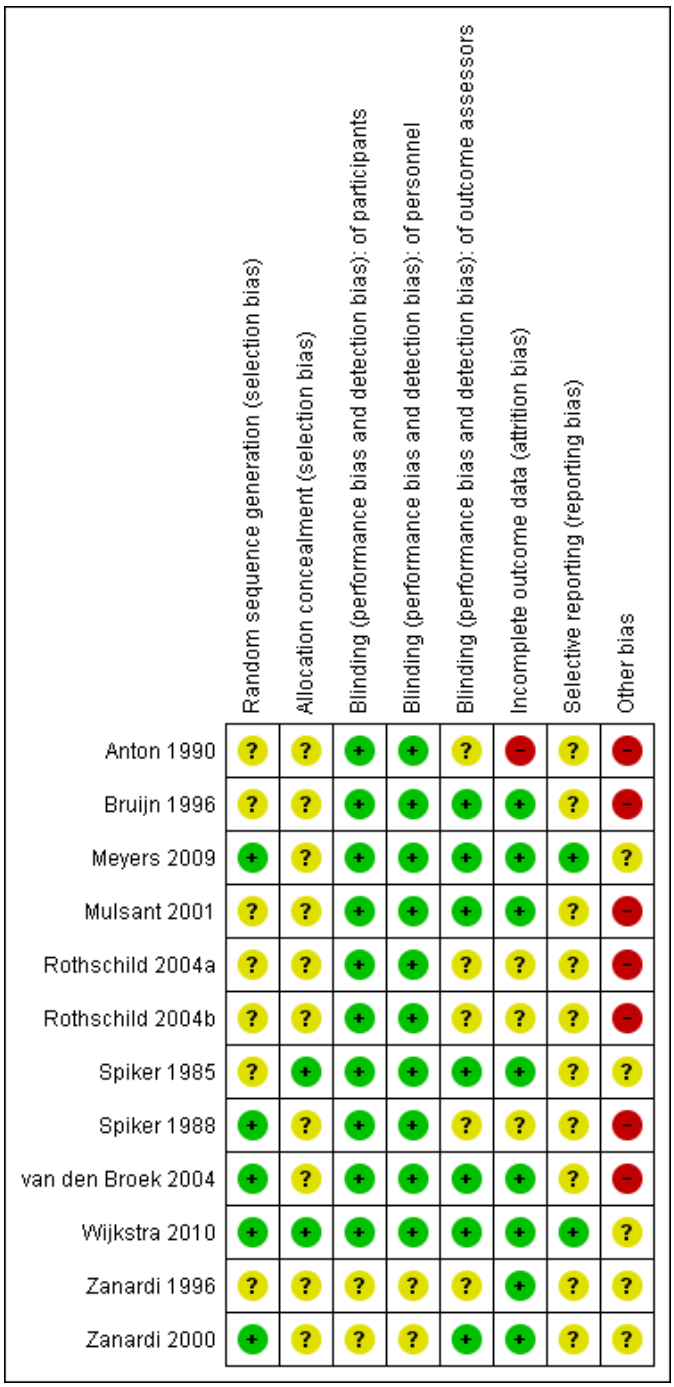


Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



If disputes arose as to which judgement should be given, resolution was achieved after consultation with the third review author (WN).

Measures of treatment effect

Binary outcomes

For binary efficacy outcomes, such as response, remission and dropouts, the risk ratio (RR) (with 95% confidence intervals (CIs)) was calculated for each comparison using the numbers randomly assigned and the numbers of events.

Continuous outcomes

For continuously distributed outcomes, we calculated the standardised mean difference (SMD).
Skewed and non-quantitative outcome data were presented descriptively.

Unit of analysis issues

We identified neither cluster-randomised nor cross-over trials. However, if found, we would not have included them in our review, as we were interested in differences not between clusters (e.g. clinics) but between drugs or classes of drugs; nor were we interested in the results of a cross-over phase, as the outcome of the first phase might have had an impact on the outcome of the second (i.e. cross-over) phase.
In case a study had multiple intervention groups, only the data for the pairwise comparison in question were included. Further, if a study compared two or more medications of the same type (e.g. venlafaxine and imipramine), data were combined into a single category, for example, the category 'antidepressant'.

Dealing with missing data

We used intention-to-treat (ITT) response data in the analyses, as ITT data are less biased and because they address a more pragmatic and clinically relevant question (Higgins 2011). When necessary, we converted response data from the trials into ITT response data by using the total number of randomly assigned participants per group that had started with treatment as the denominator. So participants who had started with medication but were withdrawn before endpoint were assumed not to have experienced response.
When data were missing, we contacted the study authors to request the required data.
No other imputation techniques were used to deal with missing data.

Assessment of heterogeneity

First, we evaluated whether clinical homogeneity could be assumed by evaluating any between-study dissimilarities regarding participants, interventions and outcome measures. Studies that were considered to threaten the clinical homogeneity assumption were excluded from the meta-analysis.
The I^2 statistic supplied with a 95% CI was used to assess the magnitude of statistical heterogeneity when values exceeding 0.40 were considered possibly relevant. We did not perform the Q test to determine heterogeneity because of its low power in our meta-analysis resulting from the low numbers of studies in all of our comparisons.
Pre-stated subgroup analyses were planned to explore sources of any heterogeneity. No meta-regression was performed.

Assessment of reporting biases

When data from at least 10 studies became available, the presence of publication bias was assessed using contour-enhanced funnel plots in which treatment effects expressed as RR (Relative Risk) from individual studies were plotted against each study's sample size. We did not perform Egger's regression test in view of low statistical power in our meta-analyses, again resulting from the low number of studies included.
It must be noted that asymmetry of funnel plots does not necessarily indicate publication bias but may result from other biases such as inflated results in smaller studies due to poorer methodological quality, or true heterogeneity.

Data synthesis

We primarily used the Mantel-Haenszel fixed-effect method to calculate pooled risk ratios with 95% confidence intervals. Risk ratios are preferred over odds ratios because of their more straightforward interpretation (i.e. the number of times the outcome is more likely to occur given one treatment over another).
In cases in which evidence of statistical heterogeneity was found, we used DerSimonian and Laird's random-effects models (DerSimonian 1986).

Subgroup analysis and investigation of heterogeneity

If sufficient data were available, subgroup analyses were planned as follows.

1. Participants who were non-refractory to prior treatment(s) during the current episode.
2. Participants with mood congruent psychotic features only.

Because all psychotically depressed patients are considered to be severely depressed, it was not considered appropriate to evaluate baseline severity in a subgroup analysis.

Sensitivity analysis

If sufficient data were available, sensitivity analyses were planned as follows.

1. Studies focusing on psychotic depression only.
2. Studies in which participants with psychotic depression were separately randomised.
3. Studies of lower methodological quality to assess robustness of results.
4. Smaller versus larger studies.

RESULTS

Description of studies

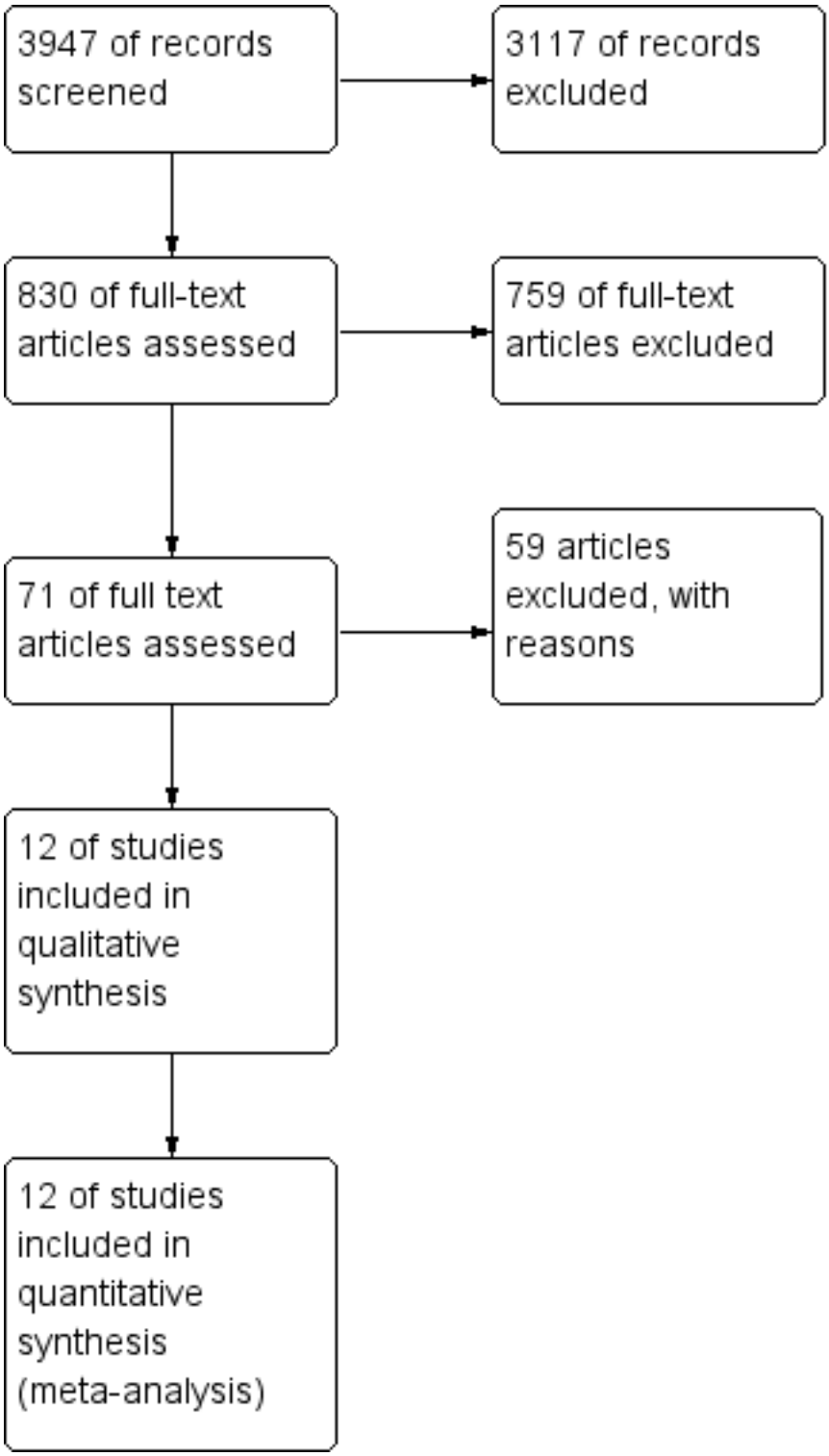
Results of the search

In 2004 from the Cochrane Central Register of Controlled Trials (CENTRAL) search, we identified 1782 publications: in the searches in MEDLINE 720 and in EMBASE 831; in total, 3333. With some overlap from 2004, we have since found an additional 326 publications. The first step of screening the abstracts of these publications in 2004 resulted in 798 relevant publications (749 from CENTRAL, 38 from MEDLINE and 11 from EMBASE); in 2011, we found 40 additional studies (three from CENTRAL, 32 from CCDANCTR-References Register and 5 from CCDANCTR-Studies Register). The second step of screening the full articles of these publications in 2004 resulted in the identification of 52 publications (47, three and two respectively) and now 17 (one from CENTRAL, 13 from CCDANCTR-References Register and three from CCDANCTR-Studies Register). In 2004, handsearching of reference lists from relevant reviews resulted in the identification of one other publication (Bellini 1994), and handsearching of the included publications did not lead to identification of any further relevant publications. Now, this resulted in no publications. In 2004, the final fourth step of reviewing these 53 publications resulted in seven included studies, and in 2004 we added two other publications, which we knew were in press: one by van den Broek (van den Broek 2004a) and one reporting two similar trials (Rothschild 2004a; Rothschild 2004b). Thus, in 2004 in total, nine publications with 10 RCTs were included. Now, in the fourth step, by reviewing 17 publications, we found two additional studies to be included.

In additional searches on the CCDANCTR and CENTRAL (December 2011 to April 2013), we found another 288 references. Of these 288 references, we did not find any other publication to include.

So, taking together both searches, we found 3974 publications through the electronic search strategies, from which we finally included 11 publications with 12 RCTs (see Figure 3).

Figure 3. Study flow diagram.



In 2004, 13 studies needed a consensus discussion in the fourth step of the search strategy before they were excluded, and now five additional studies needed such a discussion before they were excluded, resulting in a total of 18 excluded studies (see Excluded studies).

Included studies

See the table Characteristics of included studies for a description of the 12 included RCTs (Anton 1990; Bruijn 1996; Meyers 2009; Mulsant 2001; Rothschild 2004a; Rothschild 2004b; Spiker 1985; Spiker 1988; Zanardi 1996; Zanardi 2000; van den Broek 2004a; Wijkstra 2010).

Design

All studies were RCTs comparing the effects of pharmacological interventions for the treatment of psychotic depression or for the treatment of depression with and without psychotic features but with data about the participants with psychotic features published separately.

Sample sizes

Sample sizes were as follows: Anton 1990: 46; van den Broek 2004a: 48; Bruijn 1996: 30; Mulsant 2001: 54; Rothschild 2004a: 124; Rothschild 2004b: 125; Spiker 1985: 58; Spiker 1988: 27; Zanardi 1996: 32; Zanardi 2000: 22; Meyers 2009: 259; and Wijkstra 2010: 122, for a total of 947 participants,

Setting

In most studies, inpatients were included, except Meyers 2009, Rothschild 2004a and Rothschild 2004b. In Meyers 2009, 69.1% of participants entered as inpatients. In the studies of Rothschild, participants were treated for at least one week as inpatients.
Seven of the studies originate from the United States, three from the Netherlands (Bruijn 1996; van den Broek 2004; Wijkstra 2010) and two from Italy (Zanardi 1996; Zanardi 2000).

Participants

All participants fulfilled criteria for major depressive disorder with psychotic features, classified according to a formal classification system (RDC: Research Diagnostic Criteria, *DSM-III, DSM-IV, DSM-IV-TR*). Six studies explicitly used a semi-structured diagnostic interview (van den Broek 2004a: SADS: Schedule for Affective Disorders and Schizophrenia; Bruijn 1996: SADS; Mulsant 2001: BPRS: Brief Psychiatric Rating Scale; Spiker 1985: SADS; Meyers 2009: SCID IV: Structured Clinical Interview for DSM-IV; Wijkstra 2010: SCID IV).

Diagnostic procedures were different between the studies. In all studies, participants were diagnosed according to a formal classification system (RDC, *DSM-III, DSM-IV, DSM-IV-TR*). Six studies explicitly used a semi-structured diagnostic interview (van den Broek 2004a: SADS; Bruijn 1996: SADS; Mulsant 2001: BPRS; Spiker 1985: SADS; Meyers 2009: SCID IV; Wijkstra 2010: SCID IV). These different procedures could have led to possible differences in participant categories.
Seven studies included only participants with unipolar psychotic depression. In the study of Zanardi 1996, it was possible to exclude bipolar participants from the data. In the study of Anton 1990, 15.8% (six out of thirty eight) bipolar participants were included in the data that the author used in his analysis. However, it was unclear how many of the eight dropouts, who were excluded before analysis, were bipolar participants. The author was not able to give additional information. Therefore, we decided to assume a random dropout rate. In the study of Spiker 1985, 15.5% of participants were bipolar (Anton 1990; Spiker 1985). In the studies of Bruijn 1996 and Zanardi 2000, we were able to exclude the bipolar participants with additional information provided by the authors. In the studies of Bruijn 1996, Meyers 2009 and Wijkstra 2010, the types of psychotic symptoms were described in greater detail. In Spiker 1985, only participants with mood congruent delusions were included. This indicates some heterogeneity in diagnosis with regard to bipolarity.

Interventions

1. Antidepressant versus placebo was studied in one study (Spiker 1988): amitriptyline (three weeks at least 150 mg) versus placebo; treatment period was four weeks.
2. Antipsychotic versus placebo was the subject in one arm of the two identical studies of Rothschild (Rothschild 2004a; Rothschild 2004b): olanzapine (mean 11.9 mg; respectively, 14.0 mg) versus placebo; treatment period was eight weeks.
3. Antidepressant versus antidepressant was examined in five studies. In van den Broek 2004: imipramine (plasma levels imipramine plus its metabolite desmethylimipramine 192 to 521 ng/mL) versus fluvoxamine (plasma level 109 to 325 ng/mL); treatment period was four weeks after predefined blood levels were reached; in Bruijn 1996: imipramine (plasma levels imipramine plus its metabolite desmethylimipramine 199 to 400 ng/mL) versus mirtazapine (plasma level 49 to 93 ng/mL); treatment period was four weeks after predefined blood levels were reached; in Zanardi 1996: sertraline (150 mg from day eight) versus paroxetine (50 mg from day eight); treatment period was five weeks; in Zanardi 2000: venlafaxine (300 mg from day eight) versus fluvoxamine (300 mg from day eight); treatment period was five weeks; and in one arm of the study of Wijkstra 2010: imipramine (plasma levels

imipramine plus its metabolite desmethylimipramine 200 to 300 ng/mL) versus venlafaxine (375 mg); treatment period was seven weeks.

4. Antipsychotic versus antipsychotic was not available in any study.

5. Antidepressant versus antipsychotic was the subject in one arm of the study by Spiker 1985: amitriptyline (mean dose 218 mg; 130 to 500 ng/mL) versus perphenazine (mean dose 50 mg; blood level 19 to 113 ng/mL); treatment period was four weeks.

6. Antidepressant plus antipsychotic versus placebo was studied in one arm of both identical studies of Rothschild (Rothschild 2004a; Rothschild 2004b): olanzapine (12.4 mg; respectively, 13.9 mg) plus fluoxetine (23.5 mg; respectively, 22.6 mg) versus placebo; treatment period was eight weeks.

7. Antidepressant plus antipsychotic versus placebo plus antidepressant was studied in four studies. In Anton 1990: amitriptyline (150 to 250 mg) plus perphenazine (24 to 40 mg) versus amoxapine (300 to 500 mg); treatment period was four weeks; in Mulsant 2001: nortriptyline (mean 63 mg) plus perphenazine (mean 19 mg) versus nortriptyline (mean dose 76 mg); treatment started with nortriptyline, and once nortriptyline blood level was between 50 and 50 ng/mL, random assignment followed; treatment period with nortriptyline plus perphenazine (or placebo) after random assignment was 2 to 16 weeks (total treatment at least four weeks); in one arm of the study by Spiker 1985: amitriptyline (mean 170 mg; 18 to 128 ng/mL) plus perphenazine (mean 54 mg; 157 to 690 ng/mL) versus amitriptyline (mean 218 mg; 130 to 500 ng/mL); treatment period was four weeks; and in two arms of Wijkstra 2010: venlafaxine (375 mg) plus quetiapine (600 mg) versus imipramine (plasma levels imipramine plus its metabolite desmethylimipramine 200 to 300 ng/mL) and versus venlafaxine (375 mg); treatment period was 7 weeks.

8. Antidepressant plus antipsychotic versus placebo plus antipsychotic was the subject in three studies. In one arm of both identical studies of Rothschild (Rothschild 2004a; Rothschild 2004b): olanzapine (12.4 mg; respectively, 13.9 mg) plus fluoxetine (23.5 mg; respectively, 22.6 mg) versus olanzapine (mean 11.9 mg; respectively, 14.0 mg); treatment period was eight weeks; in one arm of Spiker 1985: amitriptyline (mean 170 mg; 18 to 128 ng/mL) plus perphenazine (mean 54 mg; 157 to 690 ng/mL) versus perphenazine (mean dose 50 mg; blood level 19 to 113 ng/mL); treatment period was four weeks; and in Meyers 2009, olanzapine (15 to 20 mg) plus sertraline (150 to 200 mg) versus olanzapine (15 to 20 mg); treatment period was 12 weeks.

Most of these studies had a washout period before the start of treatment or random assignment, varying from four to seven days. One study had a washout period of two weeks (Spiker 1988). Because of the design of the study of Mulsant 2001 (all participants used nortriptyline before random assignment to additional perphenazine or placebo), this study was considered a trial without a washout period. The two trials Rothschild 2004a and Rothschild 2004b had a 'screening period' of three to nine days, leaving unclear whether this was a period in which medication was not allowed. In Meyers 2009, psychotropics were tapered before random assignment without a washout period. So, heterogeneity is seen in the medication-free period before treatment.

Dosage of psychotropics used in the different trials was considered reasonably adequate. However, differences in dosing strategies led to possible bias. In four studies re TCAs, doses were adjusted according to predefined therapeutic plasma levels (Bruijn 1996; Mulsant 2001; van den Broek 2004a; Wijkstra 2010). In the study of Spiker 1985, dose adjustment was not based on plasma levels, but afterwards it was found that the plasma levels were therapeutic in most participants, although plasma levels of participants receiving TCAs alone were lower compared with those of participants receiving TCA plus antipsychotic. In the two other trials with TCAs, no plasma levels were determined. Dosages in the Spiker 1988 study were at least 150 mg/d, but only during three of the four study weeks. In the study of Anton 1990, participants received at least 150 mg/d from the third day of the four-week study period. Amitriptyline 150 mg/d is in the low range of an adequate dosage. In the studies of van den Broek 2004 and Bruijn 1996, SSRIs (fluvoxamine and mirtazapine) were dosed according to predefined plasma levels. Fixed doses were used in the studies of Zanardi (Zanardi 1996; Zanardi 2000): sertraline 150 mg, paroxetine 50 mg, venlafaxine 300 mg and fluvoxamine 300 mg. In the studies of Rothschild (Rothschild 2004a; Rothschild 2004b), doses were clinically adjusted: olanzapine 5 to 20 mg and fluoxetine 20 to 80 mg.

Differences in additional medication strategies were also noted. In most studies, additional medication was used, such as benzodiazepines (flurazepam up to 30 mg or lorazepam as needed) and anticholinergics. In the studies of van den Broek 2004a and Bruijn 1996, some participants were treated with additional haloperidol, if clinically needed. With information provided by the authors, we were able to identify these participants in the results, and we counted them as dropouts, as in these studies the focus was the comparison of two antidepressants.

Outcomes

The primary efficacy outcome was the response rate in each study. It was not possible to transfer the authors' defined response data into rates based on one definition (i.e. at least 50% reduction of the HRSD score). Some authors used response definitions based on what is often considered remission. In addition, some authors included psychotic symptoms in their response definition. In the absence of a better option, we decided to use the response data as reported by the authors.

Differences in outcome measures were noted. In most trials, the HRSD was used as an outcome measure. However, different versions of the HRSD were used, and the authors used different definitions of response. In six trials, the response definition was a reduction of at least 50% of the HRSD score compared with baseline

(Anton 1990; Bruijn 1996; Rothschild 2004a; Rothschild 2004b; van den Broek 2004; Wijkstra 2010). In four studies, the authors' definition of response was actually comparable with a frequently used definition of remission (Spiker 1985; Spiker 1988; Zanardi 1996; Zanardi 2000).

In five studies, the response definition also includes psychotic symptoms (Meyers 2009; Mulsant 2001; Spiker 1985; Zanardi 1996; Zanardi 2000). Bruijn 1996 and van den Broek 2004a used a response criterion of HRSD-17 < 50%; Anton 1990 ≤ 50%; Rothschild 2004a and Rothschild 2004b HRSD-24 < 50%; Wijkstra 2010 < 50% plus < 15; Spiker 1985 HRSD-17 < 7 and no delusions; Spiker 1988 HRSD-17 < 7 and not psychotic, or HRSD-17 6.5 to 9.5 and not psychotic and a third less of entering score; and Mulsant 2001 HRSD-17 < 11 and BPRS score of the items 11, 12 and 15 of 1 or 2 (i.e. not psychotic). In Meyers 2009, remission was defined as HRSD-17 score < 11 and no psychosis; Zanardi 1996 HRSD-21 < 8 and DDERS (Dimensions of Delusional Experience Rating Scale) = 0; and Zanardi 2000 HRSD-21 < 9 and DDERS = 0. In the studies of Zanardi 1996 and Zanardi 2000, no minimum HRSD score was applied as an inclusion criterion.

In addition to response rates based on the above criteria used by study authors, several studies reported remission rates separately (van den Broek 2004a: HRSD-17 < 8, Rothschild 2004a and Rothschild 2004b: HAMD-24 < 9 for two consecutive visits, Wijkstra 2010: HRSD-17 < 8). In two studies, the authors' definition of response is the same as what is nowadays considered the definition of remission (Spiker 1985 and Spiker 1988: HRSD-17 < 7).

Dropout rates

Overall dropout rates for the primary outcome were available for all studies. Dropout rates due to adverse effects were available for six studies (Anton 1990; Meyers 2009; Mulsant 2001; Spiker 1985; van den Broek 2004a; Wijkstra 2010). Dropout rates due to adverse effects were not based on ITT analysis for three studies (Bruijn 1996; Rothschild 2004a; Rothschild 2004b); were unavailable for one study (Spiker 1988); and were the same as the overall dropout rates for two studies (Zanardi 1996; Zanardi 2000). Dropouts specifically due to mortality or suicide were not reported, so we decided to extract only overall dropout rates.

The overall dropout rates for the studies varied from 9% to 45%. In the Bruijn 1996 study, the reported dropout rate was 10%. However, when haloperidol-treated participants were included as dropouts, as was our approach, the dropout rate was 40%. In the two multicenter trials with olanzapine/fluoxetine (Rothschild 2004a; Rothschild 2004b), dropout rates were 41%, and even higher non-completion rates (56%) were reported. Mostly, no statistically significant differences in overall dropout rates were noted between any of the treatments, neither in individual studies nor after pooling of studies.

Prior treatment

The Bruijn 1996 study provided information on prior treatment of the current episode. However, these data were not available for the subgroup with psychotic depression. In Wijkstra 2010, some data about prior treatment are available. The other studies did not provide information on prior treatment of the current episode. Therefore, it was not possible to examine differences in treatment response in relation to non-response to prior treatment(s).

Excluded studies

See the Characteristics of excluded studies table.

Reasons for exclusion involved not fulfilling inclusion criteria: open study design, problems with diagnosis (unclear diagnostic procedure, more then 20% bipolar participants included and no additional data to exclude bipolar participants), low numbers of included participants (3 to 15 per group), problems with treatments (continuation of mood stabilisers, treatment with diverse psychotropics), no possibility to extract ITT data and no adequate data on the MDD subgroup.

Risk of bias in included studies

See Figure 1 and Figure 2 and the table Characteristics of included studies.

Allocation

All included studies were randomised studies. Randomisation was fully described in van den Broek 2004a and Wijkstra 2010. Randomisation was described in part in Meyers 2009, Spiker 1985 and Spiker 1988. In the other eight studies, randomisation was mentioned as such, but the methods of randomisation were not delineated.

(Random sequence generation: seven studies unclear risk and five low risk of bias; allocation concealment: 10 unclear risk and two low risk.)

Blinding

All studies were double-blind studies, but blinding itself was not always adequately described in the methods section of the study. Blinding was adequately described in Anton 1990, Bruijn 1996, Meyers 2009, Mulsant 2001, Spiker 1985, Spiker 1988, van den Broek 2004a and Wijkstra 2010. The method of blinding was not explicitly described in Rothschild 2004a, Rothschild 2004b, Zanardi 1996 and Zanardi 2000. However, the authors explicitly state the double-blind condition of their studies; we have no reason to doubt that these double-blind conditions pertained to both investigators/assessors and participants.

(Blinding of participants: two studies unclear risk and 10 studies low risk; blinding of personnel: two studies unclear risk and 10

studies low risk; blinding of outcome assessors: five studies unclear risk and seven studies low risk.)

Incomplete outcome data

The primary efficacy outcome was the response rate (of depression). It was not possible to transfer the authors' defined response data into rates based on a single definition (i.e. 50% reduction of the HRSD score). Some authors used response definitions based on what is often considered remission (e.g. HAMD < 8 or 10), and some authors included psychotic symptoms in their response definition. In the absence of a better option, we decided to use response data as reported by the authors, with the preference to use response of depression rather than response of psychosis.

In eight of the 12 studies, we recalculated intention-to-treat response rates using all randomly assigned participants as the denominator. Of 46 participants in the study of Anton 1990, eight were dropped from the study before receiving two full weeks of active medication. These participants were excluded from the analysis by the author, but we included them in our ITT analysis. In the studies of van den Broek 2004a and Bruijn 1996, we counted participants treated with haloperidol as dropouts because haloperidol treatment in these participants was started after random assignment, in part to keep them in the study. Thus, treatment with haloperidol is considered a potential bias with regard to the effect of study medication as well as dropout. Mulsant 2001 excluded six dropouts after random assignment to perphenazine or placebo. We included them in our ITT analysis. In the studies of Rothschild 2004a and Rothschild 2004b, seven per cent and nine per cent, respectively, of randomly assigned participants were lost before baseline plus one visit. These participants were excluded from the study's analysis but were included in our ITT analysis. In both studies of Spiker (Spiker 1985; Spiker 1988), seven dropouts were excluded from their analysis, but we included them in our ITT analysis.

For the secondary outcome of change in symptom severity, we decided to refrain from using continuous data from observer depression severity scales. In two studies (Bruijn 1996; van den Broek 2004), continuous data were available for the total group (psychotic and non-psychotic) but not from the psychotic subgroup. In four studies, continuous data were not available (Anton 1990; Mulsant 2001; Spiker 1988; Zanardi 1996). In one study (Spiker 1985), baseline and final mean HRSD data are given, but it was impossible to exclude bipolar participants from these data and convert the data to ITT data. In the studies of Rothschild (Rothschild 2004a; Rothschild 2004b), only LOCF continuous data are available, and in this study, dropout is very high. Pooling of data from the three remaining studies is useless because three different comparisons are studied: AD versus AD (Zanardi 2000) with no ITT data; AD + AP versus AP (Meyers 2009); and AD + AP versus AD (Wijkstra 2010).

(Incomplete data: one study high risk, three studies unclear risk and eight studies low risk.)

Selective reporting

All studies used generally accepted outcomes. From two recent studies (Meyers 2009 and Wijkstra 2010), the study protocol is available; in these two studies, no post-protocol changes in outcome measures had been made (except in Wijkstra 2010, which used remission as an outcome not stated in the protocol but with reasonable argument that remission has become a generally accepted outcome measure).

(Selective reporting: 10 studies unclear risk and two studies low risk.)

Other potential sources of bias

The subgroup of psychotic depressed participants (Bruijn 1996; Spiker 1988; van den Broek 2004a) was not stratified before random assignment in any of these studies. Subgroup analysis are more likely to be carried out if the results for primary outcomes are not significant or are more likely to be reported for groups for whom a significant result was found. However, in the case of this review, we ourselves analysed subgroup data in studies that primarily reported on participants with depression with and without psychotic features.

As described under 'Description of included studies', clinical heterogeneity is seen within the results.

In five studies, response definition included response with regard to psychotic symptoms (Meyers 2009; Mulsant 2001; Spiker 1985; Zanardi 1996; Zanardi 2000), and in the other studies (Anton 1990; Bruijn 1996; Rothschild 2004a; Rothschild 2004b; Spiker 1988; van den Broek 2004; Wijkstra 2010; Zanardi 1996), response rates concerned only change in severity of depression. In our analysis, we looked only for response of depression leading to a possible bias favouring antidepressants over antipsychotics.

(Other biases together: seven studies high risk and five studies unclear risk.)

Effects of interventions

For the secondary outcome of change in symptom severity, extracting continuous data from observer depression severity scales was not possible because in seven studies, we were not able to convert these data according to intention-to-treat analysis (Anton 1990; Bruijn 1996 ; Rothschild 2004a; Rothschild 2004b; Spiker 1985; Spiker 1988; van den Broek 2004a), and in two other studies, no continuous data were reported (Zanardi 1996; Zanardi 2000).

Comparison 1. Antidepressant versus placebo

Primary outcomes

1.1 Efficacy response rates

An antidepressant was compared with placebo in only one study. In this study (Spiker 1988; Analysis 1.1), no evidence suggested that amitriptyline was superior to placebo (RR 8.40, 95% CI 0.50 to 142.27, P = 0.14).

1.2 Harm: overall dropout rate during acute treatment

No difference (Spiker 1988; Analysis 1.2) (RR 1.24, 95% CI 0.34 to 4.51).

Secondary outcomes

1.3 Remission

No data.

1.4 Change from baseline

No ITT data.

1.5 Quality of life

No data,

1.6 Dropout rate due to adverse effects

No data.

Comparison 2. Antipsychotic versus placebo

Primary outcomes

2.1 Efficacy response rates

Two studies compared an antipsychotic with placebo (Rothschild 2004a; Rothschild 2004b). Pooling of these studies (Analysis 2.1) does not reveal a difference between olanzapine and placebo (RR 1.13, 95% CI 0.74 to 1.73, P = 0.57).

2.2 Harm: overall dropout rate during acute treatment

Pooling of these two studies (Rothschild 2004a; Rothschild 2004b; Analysis 2.2) reveals no difference (RR 0.79, 95% CI 0.57 to 1.08, P = 0.14).

Secondary outcomes

2.3 Remission

No data.

2.4 Change from baseline

For the secondary outcome of change in symptom severity, extracting continuous data from observer depression severity scales was not possible; we were not able to convert these data according to ITT analysis (Rothschild 2004a; Rothschild 2004b).

2.5 Quality of life

No data.

2.6 Dropout rate due to adverse effects

No data.

Comparison 3. Antidepressant versus antidepressant

Primary outcomes

3.1 Efficacy response rates

In five studies, two different antidepressants were compared directly with each other (Bruijn 1996; van den Broek 2004a; Wijkstra 2010; Zanardi 1996; Zanardi 2000).

Differences in efficacy were found in three of these studies. In the study by van den Broek 2004a (Analysis 3.1), imipramine was more effective than fluvoxamine (RR 2.10, 95% CI 1.06 to 4.17); in the study by Bruijn 1996 (Analysis 3.1), imipramine was near more effective than mirtazapine (RR 3.00, 95% CI 1.01 to 8.95); and in the study by Zanardi 1996 (Analysis 3.1), sertraline was better than paroxetine (RR 3.37, 95% CI 1.19 to 9.57, P = 0.02). The study by Zanardi 2000 (Analysis 3.1) did not reveal a difference between fluvoxamine and venlafaxine (RR 1.50, 95% CI 0.82 to 2.75). In Wijkstra 2010 (Analysis 3.1), no difference was noted between imipramine and venlafaxine (RR 1.57, 95% CI 0.93 to 2.67).

Pooling of studies was not reasonable because different antidepressants were used.

3.2 Harm: overall dropout rate during acute treatment

In these five studies, no differences in overall dropout rate were reported (Analysis 3.1).

Secondary outcomes

3.3 Remission

No data.

3.4 Change from baseline

Pooling not possible because different antidepressants were used.

3.5 Quality of life

No data.

3.6 Dropout rate due to adverse effects

No data.

Comparison 4. Antipsychotic versus antipsychotic

No data were available for any outcome for this comparison.

Comparison 5. Antidepressant versus antipsychotic

Primary outcomes

5.1 Efficacy response rates

We found one RCT that compared an antidepressant with an antipsychotic. In this trial (Spiker 1985; Analysis 4.1), no difference between perphenazine and amitriptyline was reported (RR 2.09, 95% CI 0.64 to 6.82).

5.2 Harm: overall dropout rate during acute treatment

One study (Spiker 1985; Analysis 4.2) reported no difference between perphenazine and amitriptyline (RR 1.79, 95% CI 0.18 to 18.02).

Secondary outcomes

5.3 Remission

No data.

5.4 Change from baseline

For the secondary outcome of change in symptom severity, extracting continuous data from observer depression severity scales was not possible because in the only study, we were not able to convert these data according to intention-to -treat analysis (Spiker 1985).

5.5 Quality of life

No data.

5.6 Dropout rate due to adverse effects

No data.

Comparison 6. Antidepressant plus antipsychotic versus placebo

Primary outcomes

6.1 Efficacy response rates

Two identical studies (Rothschild 2004a; Rothschild 2004b) compared the combination of fluoxetine and olanzapine with placebo. Pooling of the studies (Analysis 5.1) reveals the combination to more efficacious (RR 1.86, 95% CI 1.23 to 2.82, P = 0.003).

6.2 Harm: overall dropout rate during acute treatment

Pooling of both studies (Rothschild 2004a; Rothschild 2004b; Analysis 5.2) reveals no difference (RR 0.75, 95% CI 0.48 to 1.18, P = 0.21).

In this pooling, we find heterogeneity (I^2 = 76%). In these two studies, dropout rates are high-in one study higher for placebo, and in the other study higher for olanzapine.

Secondary outcomes

6.3 Remission

No data.

6.4 Change from baseline

For the secondary outcome of change in symptom severity, extracting continuous data from observer depression severity scales was not possible because in these two studies, we were not able to convert these data according to intention-to -treat analysis (Rothschild 2004a; Rothschild 2004b).

6.5 Quality of life

No data.

6.6 Dropout rate due to adverse effects

No data.

Comparison 7. Antidepressant plus antipsychotic versus placebo plus antipsychotic

Primary outcomes

7.1 Efficacy response rates

Four studies compared the combination of an antidepressant plus an antipsychotic with antipsychotic monotherapy (Meyers 2009; Spiker 1985; Rothschild 2004a; Rothschild 2004b).
In the ITT analysis of Spiker et al. (Spiker 1985; Analysis 6.1), the combination of amitriptyline plus perphenazine was superior to perphenazine (RR 3.61, 95% CI 1.23 to 10.56, P = 0.02).
Pooling of both identical Rothschild studies (Analysis 6.1; Rothschild 2004a; Rothschild 2004b) gives an advantage to the combination of olanzapine plus fluoxetine over olanzapine alone (RR 1.64, 95% CI 1.10 to 2.44, P = 0.01).
In Meyers at al (Meyers 2009; Analysis 6.1), olanzapine plus sertraline was more effective than olanzapine alone (RR 0.57, 95% CI 0.39 to 0.82, P = 0.003).
Pooling the data from all four studies (Meyers 2009; Rothschild 2004a; Rothschild 2004b; Spiker 1985; Analysis 6.1) shows a difference favouring the combination (RR 1.83, 95% CI 1.40 to 2.38, P = 0.00001).

7.2 Harm: overall dropout rate during acute treatment

Pooling all four studies (Meyers 2009; Rothschild 2004a; Rothschild 2004b; Spiker 1985; Analysis 6.2) reveals no difference (RR 0.79, 95% CI 0.63 to 1.01, P = 0.06, I^2 = 63%).
Heterogeneity of these pooled results is possibly a result of the Rothschild 2004a and Rothschild 2004b studies because in these two studies, dropout rates are very high, and these rates are different between the two studies.

Secondary outcomes

7.3 Remission

No data.

7.4 Change from baseline

For the secondary outcome of change in symptom severity, extracting continuous data from observer depression severity scales was not possible because in three studies, we were not able to convert these data according to ITT analysis (Rothschild 2004a; Rothschild 2004b; Spiker 1985).

7.5 Quality of Life

No data.

7.6 Dropout rate due to adverse effects

No data.

Comparison 8. Antidepressant plus antipsychotic versus placebo plus antidepressant

Primary outcomes

8.1 Efficacy response rates

Four studies (five comparisons) compared the combination of an antidepressant plus an antipsychotic with antidepressant monotherapy (Anton 1990; Mulsant 2001; Spiker 1985; Wijkstra 2010). Two of these studies compared a TCA plus an antipsychotic with TCA monotherapy (Mulsant 2001; Spiker 1985). Pooling of the ITT data of these two studies (as in our first review) did not reveal a difference (RR 1.44, 95% CI 0.86 to 2.41, P = 0.16: analysis in first review). Another study that compared the combination of venlafaxine plus an antipsychotic (quetiapine) with venlafaxine monotherapy and with imipramine monotherapy (Wijkstra 2010; Analysis 7.1) found a difference with venlafaxine (RR 0.51, 95% CI 0.31 to 0.83) but not with imipramine (RR 0.80, 95% CI 0.55 to 1.14). In the study of Anton 1990, amoxapine was used as an antidepressant with antipsychotic effects. Pooling of these five studies results in a difference favouring the combination (RR 1.42; 95% CI 1.11 to 1.80, P = 0.002; Analysis 7.1).
Three studies in this comparison used the same antidepressant in both arms, whereas two studies used a different antidepressant in each arm. Analysing just the studies where the same antidepressant was used in both arms (i.e. leaving out Anton 1990 and Wijkstra 2010) (Analysis 8.1) still reveals a difference in favour of the combination (RR 1.70, 95% CI 1.19 to 12.43, P = 0.003).

8.2 Harm: overall dropout rate during acute treatment

No differences were noted after pooling of these three studies (Analysis 7.2): RR 0.91, 95% CI 0.55 to 1.50, P = 0.69. In the subgroup analysis, no difference is noted (Analysis 8.2; RR 1.04, 95% CI 0.52 to 0.91, P = 0.91).

Secondary outcomes

8.3 Remission

No data.

8.4 Change from baseline

For the secondary outcome of change in symptom severity, extracting continuous data from observer depression severity scales was not possible because we were not able to convert these data according to intention-to-treat analysis.

8.5 Quality of life

No data.

8.6 Dropout rate due to adverse effects

No data.

Subgroup and sensitivity analysis

Because of lack of data, subgroup and sensitivity analyses were not possible.

DISCUSSION

Despite our extensive search of the literature (screening more than 3600 abstracts and 829 full articles, as well as carefully reviewing 69 publications), we identified very few RCTs on pharmacological treatment with an antidepressant, an antipsychotic or the combination of an antidepressant with an AP of participants with a major depressive disorder with a current episode with psychotic features (unipolar psychotic depression).

In addition to nine studies, whose main focus was the treatment of participants with psychotic depression, we were able to find three studies that reported separately on the effects of the subgroups of participants with psychotic depression. The authors of two further studies provided us with additional information on the results for the subgroup of psychotically depressed participants in their studies on both psychotic and non-psychotic depressed patients. In our previous review, we invited authors of several studies to provide us with subgroup data if available, so we could use these data in our review, but we did not receive any data.

In this update, we found two new studies to be added to the ten studies found in our previous review, which (again) illustrates that this most severe form of depression is highly under-investigated. One probable reason for this is that it is very difficult to conduct RCTs in patients with psychotic depression. These patients often are not only psychotic but are very anxious or physically ill. Moreover, they are often offered ECT directly without a trial of pharmacological treatment. Finally, many patients with psychotic depression are not able or are reluctant to give informed consent, or they tend to withdraw their consent.

As a result of the paucity of RCTs, few clinically relevant conclusions can be drawn.

Summary of main results

1. Evidence was derived from two identical placebo-controlled studies (Rothschild 2004a; Rothschild 2004b; Analysis 5.1) for the efficacy of the combination of an antidepressant plus an antipsychotic (fluoxetine plus olanzapine vs placebo; RR 1.86, 95% CI 1.23 to 2.82, P = 0.003)
2. We found evidence that the combination of an antidepressant plus an antipsychotic is more effective than antidepressant monotherapy. Pooling of four studies (five comparisons) that compared the combination of an antidepressant plus an antipsychotic with antidepressant monotherapy (Anton 1990; Mulsant 2001; Spiker 1988; Wijkstra 2010; Analysis 7.1) results in a significant difference favouring the combination (RR 1.42, 95% CI 1.11 to 1.80, P = 0.002). When the two comparisons with a different antidepressant are left out (Anton 1990; Wijkstra 2010), this difference still is statistically significant (Analysis 8.1; RR 1.70, 95% CI 1.19 to 2.43, P = 0.003). In our previous review, based on two studies (Mulsant 2001; Spiker 1988; RR 1.44, 95% CI 0.86 to 2.41), no such proof was found. Therefore, it can be concluded that the recommendation in the APA guidelines (APA 2010) that in psychotic depression the combination should be preferred over an AD alone can be considered evidence based.
3. The combination of an antidepressant plus an antipsychotic is more effective than antipsychotic monotherapy. Pooling of four studies that compared the combination of an antidepressant plus an antipsychotic with antipsychotic monotherapy (Meyers 2009; Rothschild 2004a; Rothschild 2004b; Spiker 1985; Analysis 6.1) shows a difference favouring the combination (RR 1.83, 95% CI 1.40 to 2.38, P = 0.00001). This confirms the conclusion stated in our previous review based on three of these four studies (Rothschild 2004a; Rothschild 2004b; Spiker 1985).
4. No randomised controlled data are available to lead to the conclusion that an antidepressant alone is efficacious in the treatment of psychotic depression. Only one small study compared monotherapy with an antidepressant (amitriptyline) with placebo (Spiker 1988; Analysis 1.1) and reported no difference (RR 8.40, 95% CI 0.50 to 142.27).
5. No randomised controlled data are available to lead to the conclusion that an antipsychotic alone is efficacious in the treatment of psychotic depression. Two studies compared monotherapy with an antipsychotic (olanzapine) with placebo (Rothschild 2004a;

Rothschild 2004b). Pooling of these two studies does not reveal a difference (Analysis 2.1; RR 1.13, 95% CI 0.74 to 1.73).
6. We were not able to collect data on prior treatments. So we could not address the second objective of our review: to assess whether differences in response to treatment in the current episode would be related to non-response to prior treatment(s).
7. Regarding acceptability of treatment (dropout, adverse effects) and quality of life, we were able to collect data only on overall dropout. In all studies except the study of Meyers 2009, no differences in overall dropout rates were reported between any of the treatment groups, neither in individual studies nor after pooling of studies. With this rather rough measure, we did not find overall differences in overall acceptability of treatments. In the study of Meyers 2009, fewer dropouts were reported in the group treated with the combination of olanzapine plus sertraline than in the group treated with olanzapine alone (RR 1.43, 95% CI 1.08 to 1.88). The authors suggest that higher attrition rates among participants treated with olanzapine plus sertraline may be attributable to insufficient response in the olanzapine-treated group.
8. We only found indication of heterogeneity drop-out data in instances where the two identical studies of Rothschild (Rothschild 2004a; Rothschild 2004b) were included (Analysis 2.2; Analysis 5.2; Analysis 6.2). This heterogeneity probably is due to high dropout rates, together with differences in dropout rates between the two studies.

Overall completeness and applicability of evidence

Our conclusions are based on only 12 studies that fulfilled our inclusion criteria. Moreover, in the included studies, only a few different antidepressants and antipsychotics were used. Therefore, it remains unclear whether the above conclusions can be extrapolated to other antidepressants and antipsychotics. Strictly spoken, evidence that the combination of an antidepressant plus an antipsychotic is more effective than an antidepressant alone in an RCT has been obtained only for the combination of venlafaxine and quetiapine compared with venlafaxine (Wijkstra 2010), although evidence that the combination of an antidepressant plus an antipsychotic is more effective than an antipsychotic alone has been obtained in RCTs for only two antipsychotics: the combination of amitriptyline and perphenazine reported as more effective than perphenazine (Spiker 1985); all three other studies involved combinations with olanzapine: in two studies with fluoxetine (Rothschild 2004a; Rothschild 2004b) and in one study with sertraline (Meyers 2009).
Nearly all participants in these studies were inpatients. This of course is a consequence of the severity of the illness and the fact that in clinical practice, most patients with psychotic depression are treated as inpatients. There is also the problem of patients who were not included in the studies: the most severely ill patients who were not able to give informed consent or who were immediately given ECT. All these points limit the generalisability of the results to all patients with psychotic depression.
These problems with the quality of diagnostic assessment can restrict generalisation of the findings to all patients with psychotic depression, leaving aside the problem of establishing the diagnosis of psychotic depression in clinical reality.

Quality of the evidence

The strength of this review and its conclusion is that only randomised controlled studies have been included, and only ITT data have been used in the analyses.
Several factors limit our confidence in the findings of this review. Most studies were relatively small. Only four RCTs had a more or less adequate sample size: Rothschild 2004a; Rothschild 2004b: with olanzapine 48 and 53 participants, with placebo 51 and 49 participants, but only 25 and 23 olanzapine plus fluoxetine, respectively; Meyers 2009: with olanzapine + sertraline 129 and sertraline 130 participants; and Wijkstra 2010: imipramine 42, venlafaxine 39 and venlafaxine + quetiapine 41 participants. As with all systematic reviews, publication bias is a potentially serious source of bias. Too few studies were identified to allow further investigation into the possibility of publication bias (e.g. by making funnel plots). However, the fact that in these mostly small trials, five (50%) found a significant difference between two active treatments, is suggestive of publication bias.
Allocation concealment, especially in the older studies, was not explicitly described. Although we do not assume allocation concealment to be a real bias, this is of course unsure.
Dosages of antidepressants and antipsychotics used in the different trials were considered reasonably adequate. However, differences in dosing strategies were noted, leading to possible bias. Differences in additional medication strategies and differences in treatment periods were also reported (see paragraph 'Other potential sources of bias').
We could use only one outcome measure regarding efficacy: the response rates as defined by the authors. It was impossible to recalculate these response rates into a standard response rate based on a single definition (e.g. reduction on the HRSD-17 of at least 50% compared with baseline), as many studies used other versions of the HRSD. Moreover, some studies used response definitions that are commonly used for the definition of remission, and we could not recalculate the results to the more commonly used definition of response (i.e. reduction of 50% compared with baseline).
We cannot rule out the possibility of differences in the diagnostic assessment of participants and thus in the quality of the diagnosis across the studies. Although it was reported in all publications that participants included in the trials fulfilled criteria for a major depressive episode with psychotic features, according to a specified diagnostic classification system (RDC, *DSM-III* or *DSM-IV*), one could doubt the reliability of the diagnoses made in some trials. Six studies (Bruijn 1996; Meyers 2009; Mulsant 2001; Spiker

1985; van den Broek 2004; Wijkstra 2010) used a semi-structured interview, and only three studies (Bruijn 1996; Meyers 2009; Wijkstra 2010) reported the types of psychotic features. This leaves open the possibility that for instance the judgement that a feeling or idea of guilt was a guilt delusion was drawn differently across the studies in this review. A similar diagnostic problem may have played a role in the judgement of whether a participant had a psychotic depression in the course of unipolar disorder or bipolar disorder. Finally, the study of Mulsant 2001 included a geriatric sample with a mean age of 72 years, leading to the possibility that dementia or another neurological disorder was part of the diagnosis in some participants.

Potential biases in the review process

One problem is that there does not exist a key word (Mesh Term) for psychotic depression. Therefore, we had to search all RCTs involving depression on whether included participants had depression with psychotic features, or whether such participants had been part of the group of included participants and were reported as a separate subgroup. We anticipated in the first version of this review that we might have missed one or more studies. However, we did not receive any information that we had missed any study. Therefore, in this update, we are now rather sure that we indeed have included all existing studies.

In three studies (Bruijn 1996; Spiker 1988; van den Broek 2004a), the subgroup of psychotic depressed participants was part of a greater group of participants with psychotic and non-psychotic depression, although the subgroups were not stratified before random assignment.

Another potential problem, which was not taken into account in our a priori protocol before this systematic review was performed, is that in five studies, the response definition included response with regard to psychotic symptoms (Meyers 2009; Mulsant 2001; Spiker 1985; Zanardi 1996; Zanardi 2000), although in the other studies (Anton 1990; Bruijn 1996; Rothschild 2004a; Rothschild 2004b; Spiker 1988; van den Broek 2004; Wijkstra 2010; Zanardi 1996), response rates concerned only change in severity of depression. In our analysis, we looked only for the response of depression, leading to a possible bias favouring antidepressants over antipsychotics.

Agreements and disagreements with other studies or reviews

In a review of practice guidelines regarding the treatment of psychotic depression (Wijkstra 2007), we found different recommendations based on slightly different studies, mostly not reanalyzed. Two guidelines-Nice 2004 and the Dutch guideline 2005 (now updated with no changes regarding treatment of psychotic depression: Dutch Guideline 2009; NICE 2009)-were cautious in their recommendations. Nice 2004: Augmenting an AD with an AP should be considered; and Dutch 2005: Starting treatment with a TCA and if after 4 weeks there is still no response, adding an AP is a reasonable option; starting with the combination of a TCA and an AP is also a reasonable option. The other reviewed guidelines, including the APA guideline 2000 (now updated with no differences regarding the treatment of psychotic depression; APA 2010), recommend using the combination of an AD and an AP. None of these guidelines recommended monotherapy with an antipsychotic. These recommendations were not based on a systematic review of data from all available RCTs; they were based on a few studies-some randomised and some open non-randomised. Another meta-analysis on the treatment of psychotic depression has been published (Farahani 2012). This review and meta-analysis is focused on the comparison of antidepressant or antipsychotic monotherapy with combination treatment. Five studies were included (Anton 1990; Künzel 2008; Mulsant 2001; Spiker 1985; Wijkstra 2010). In our review, for this particular comparison, we excluded Künzel 2008 because it is unclear whether less or more than 20% of included participants in the ITT group had a bipolar disorder (see table Excluded studies), and we did not use the study of Anton 1990 because we did not regard amoxapine as an antidepressant (Analysis 8.1). We included in our review the same three other studies (Mulsant 2001; Spiker 1985; Wijkstra 2010), using exactly the same extracted ITT data. The conclusion of this review is consistent with ours: Combination treatment is more effective than antidepressant monotherapy. For the comparison of an antipsychotic plus an antidepressant versus an antipsychotic, the same four studies with again exactly the same extracted ITT data were included (Meyers 2009; Rothschild 2004a; Rothschild 2004b; Spiker 1985; Analysis 5.1), leading in both reviews to the same conclusion: Combination treatment is more effective than antipsychotic monotherapy. As in our review, no differences were reported in overall dropout rates across all studies for both comparisons.

AUTHORS' CONCLUSIONS

Implications for practice

Psychotic depression is heavily under-studied, limiting confidence in the conclusions drawn. Evidence suggests that combination therapy with an antidepressant plus an antipsychotic is more effective than either treatment alone or placebo. Evidence for treatment with an antidepressant alone or for an antipsychotic alone is lacking.

Implications for research

Further studies are needed:

1. to study the efficacy of other combinations of an antidepressant plus an antipsychotic. Regarding the antidepressants: combinations with a TCA, with an SSRI or with other newer antidepressants such as mirtazapine; regarding the antipsychotics: combinations with other so called atypical antipsychotics (aripiprazole, risperidone, olanzapine, etc.);

2. to compare the effect of the combination of an antidepressant with antipsychotics with other pharmacological options, such as the augmentation of an antidepressant with lithium or more experimental treatments such as ketamine;

3. to compare the effect of the combination of an antidepressant plus an antipsychotic with ECT; and

4. to evaluate the efficacy of stepwise approaches or algorithms encompassing the previous steps after each other.

ACKNOWLEDGEMENTS

We would like to thank Dr J. Bruijn, Dr W.W. van den Broek and Dr R. Zanardi for providing additional information about their studies. A C ipriani is supported by the NIHR Oxford Cognitive Health Clinical Research Facility and was expert witness for Accord Healthcare for a patent issue about quetiapine extended release.

CRG Funding Acknowledgement:

The National Institute for Health Research (NIHR) is the largest single funder of the Cochrane Depression, Anxiety and Neurosis Group.

Disclaimer:

The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the NIHR, NHS or the Department of Health.

REFERENCES

References to studies included in this review

Anton 1990 *{published data only}*

* Anton RF Jr, Burch EA Jr. Amoxapine versus amitriptyline combined with perphenazine in the treatment of psychotic depression. *American Journal of Psychiatry* 1990;**147**(9): 1203–8.

Anton RF Jr, Burch EA Jr. Response of psychotic depression subtypes to pharmacotherapy. *Journal of Affective Disorders* 1993;**28**(2):125–31.

Bruijn 1996 *{published data only}*

Bruijn JA, Moleman P, Mulder PG, van den Broek WW. Depressed in-patients respond differently to imipramine and mirtazapine. *Pharmacopsychiatry* 1999;**32**(3):87–92.

* Bruijn JA, Moleman P, Mulder PG, van den Broek WW, van Hulst AM, van der Mast RC, et al. A double-blind, fixed blood-level study comparing mirtazapine with imipramine in depressed in-patients. *Psychopharmacology* 1996;**127**(3):231–7.

Meyers 2009 *{published data only}*

Andreescu C, Mulsant BH, Peasley-Miklus C, Rothschild AJ, Flint AJ, Heo M, et al. Persisting low use of antipsychotics in the treatment of major depressive disorder with psychotic features. Journal of Clinical Psychiatry 2007; Vol. 68, issue 2:194–200.

Blumberger DM, Mulsant BH, Emeremni C, Houck P, Andreescu C, Mazumdar S, et al. Impact of prior pharmacotherapy on remission of psychotic depression in a randomized controlled trial. Journal of Psychiatric Research 2011; Vol. 45, issue 7:896–901.

Flint AJ. Challenges in conducting a pharmacotherapy study in psychotic depression. *157th Annual Meeting of the American Psychiatric Association; 2004 May 1-6; New York, NY* 2004:No. 63D.

Meyers B, Rothschild A, Gabriele M, English J, McShea M, Kamara T, et al. Effectiveness of Selective Serotonin Reuptake Inhibitors Combined With Antipsychotic Medication for the Treatment of Psychotic Depression [STOP-PD; NCT00056472]. ClinicalTrials.gov [http://www.clinicaltrials.gov] 2004.

Meyers BS. Diagnosis and treatment of psychosis in geriatric major depression. *46th Annual NCDEU (New Clinical Drug Evaluation Unit) Meeting; 2006 June 12 - 15; Boca Raton, FL* 2006:83.

* Meyers BS, Flint AJ, Rothschild AJ, Mulsant BH, Whyte EM, Peasley-Miklus C, et al. A double-blind randomized controlled trial of olanzapine plus sertraline vs olanzapine plus placebo for psychotic depression: the study of pharmacotherapy of psychotic depression (STOP-PD). Archives of General Psychiatry 2009; Vol. 66, issue 8: 838–47.

Meyers BS, Flint AJ, Rothschild AJ, Mulsant BH, Whyte EM, Peasley-Miklus C, et al. Errors in Table in: A Double-blind Randomized Controlled Trial of Olanzapine Plus Sertraline vs Olanzapine Plus Placebo for Psychotic Depression: the Study of Pharmacotherapy of Psychotic Depression (STOP-PD). *Archives of General Psychiatry* 2011;**68**(6):625–6.

Schaffer A, Flint AJ, Smith E, Rothschild AJ, Mulsant BH, Szanto K, et al. Correlates of suicidality among patients with psychotic depression. *Suicide and Life-Threatening Behavior* 2008;**38**(4):403–14.

Weissman J, Flint A, Meyers B, Ghosh S, Mulsant B, Rothschild A, et al. Factors associated with non-completion in a double-blind randomized controlled trial of olanzapine plus sertraline versus olanzapine plus placebo for psychotic depression. *Psychiatry Research* 2012;**197**(3):221–6.

Mulsant 2001 *{published data only}*
Mulsant BH, Sweet RA, Rosen J, Pollock BG, Zubenko GS, Flynn T, et al. A double-blind randomized comparison of nortriptyline plus perphenazine versus nortriptyline plus placebo in the treatment of psychotic depression in late life. *Journal of Clinical Psychiatry* 2001;**62**(8):597–604.

Rothschild 2004a *{published data only}*
* Rothschild AJ, Williamson DJ, Tohen MF, Schatzberg A, Andersen SW, Van Campen LE, et al. A double-blind, randomized study of olanzapine and olanzapine/fluoxetine combination for major depression with psychotic features. Journal of Clinical Psychopharmacology 2004; Vol. 24, issue 4:365–73.
Williamson DJ, Andersen SW, Campen LE, Sanger TM, Paul S, Corya SA, et al. Olanzapine-fluoxetine combination for difficult-to-treat depression. *154th Annual Meeting of the American Psychiatric Association; 2001 May 5-10; New Orleans; LA* 2001:NR549.

Rothschild 2004b *{published data only}*
Rothschild AJ, Williamson DJ, Tohen MF, Schatzberg A, Andersen SW, Van Campen LE, et al. A double-blind, randomized study of olanzapine and olanzapine/fluoxetine combination for major depression with psychotic features. *Journal of Clinical Psychopharmacology* 2004;**24**(4):365–73.
Williamson DJ, Andersen SW, Campen LE, Sanger TM, Paul S, et al. Olanzapine-fluoxetine combination for difficult-to-treat depression. *154th Annual Meeting of the American Psychiatric Association; 2001 May 5-10; New Orleans; LA* 2001:NR549.

Spiker 1985 *{published data only}*
Spiker D, Perel JM, Hanin I, Dealy RS, Griffin SJ, Soloff PH, et al. The pharmacological treatment of delusional depression: part II. *Journal of Clinical Psychopharmacology* 1986;**6**(6):339–42.
Spiker DG, Hanin I, Perel JM, Cofsky JE, Rossi AJ, Sorisio D. Pharmacological treatment of delusional depressives. *Psychopharmacology bulletin* 1982;**18**(4):184–6.
* Spiker DG, Weiss JC, Dealy RS, Griffin SJ, Hanin I, Neil JF, et al. The pharmacological treatment of delusional depression. *American Journal of Psychiatry* 1985;**142**(4): 430–6.

Spiker 1988 *{published data only}*
Spiker DG, Kupfer DJ. Placebo response rates in psychotic and nonpsychotic depression. *Journal of Affective Disorders* 1988;**14**(1):21–3.

van den Broek 2004 *{published data only}*
van den Broek WW, Birkenhäger TK, Mulder PG, Bruijn J, Moleman P. A double blind study comparing imipramine with fluvoxamine in depressed inpatients. Psychopharmacology 2004; Vol. 175, issue 4:481–6.

Wijkstra 2010 *{published data only}*
Wijkstra J, Burger H, van den Broek WW, Birkenhäger TK, Janzing JG, Boks MP, et al. Treatment of unipolar psychotic depression: a randomized, double-blind study comparing imipramine, venlafaxine, and venlafaxine plus quetiapine. *Acta Psychiatrica Scandinavica* 2010;**121**(3):190–200.

Zanardi 1996 *{published data only}*
Zanardi R, Franchini L, Gasperini M, Perez J, Smeraldi E. Double-blind controlled trial of sertraline versus paroxetine in the treatment of delusional depression. *American Journal of Psychiatry.* 1996;**153**(12):1631–3.

Zanardi 2000 *{published and unpublished data}*
Zanardi R, Franchini L, Serretti A, Perez J, Smeraldi E. Venlafaxine versus fluvoxamine in the treatment of delusional depression: a pilot double-blind controlled study. *Journal of Clinical Psychiatry* 2000;**61**(1):26–9.

References to studies excluded from this review

Belanoff 2001 *{published data only}*
Belanoff JK, Flores BH, Kalezhan M, Sund B, Schatzberg AF. Rapid reversal of psychotic depression using mifepristone. *Journal of Clinical Psychopharmacology* 2001;**21**(5):516–21.

Bellini 1994 *{published data only}*
Bellini L, Gasperini M, Gatti F, Franchini L, Smeraldi E. A double blind study with fluvoxamine vs. desipramine combined with placebo or haloperidol in delusional depression. *Critical Issues in the Treatment of Affective Disorders* 1994;**9**:32–6.

Blasey 2009 *{published data only}*
Blasey CM, DeBattista C, Roe R, Block T, Belanoff JK. Corrigendum to 'A Multisite Trial of Mifepristone for the Treatment of Psychotic Depression: A Site-by-Treatment Interaction' [Contemp. Clin. Trials 30 (2009) 284-288] (DOI:10.1016/j.cct.2009.03.001). *Contemporary clinical trials* 2010;**31**(1):134.
Blasey CM, DeBattista C, Roe R, Block T, Belanoff JK. Corrigendum to 'A Multisite Trial of Mifepristone for the Treatment of Psychotic Depression: A Site-by-Treatment Interaction' [Contemp. Clin. Trials 30 (2009)284-288] (DOI:10.1016/j.cct.2009.03.001). *Contemporary clinical trials* 2009;**30**(5):497.
* Blasey CM, Debattista C, Roe R, Block T, Belanoff JK. A multisite trial of mifepristone for the treatment of psychotic depression: a site-by-treatment interaction. *Contemporary Clinical Trials* 2009;**30**(4):284–8.

Blasey 2011 *{published data only}*
Blasey CM, Block TS, Belanoff JK, Roe RL. Efficacy and safety of mifepristone for the treatment of psychotic depression. *Journal of Clinical Psychopharmacology* 2011;**31** (4):436–440.

Cassacchia 1984 *{published data only}*
Casacchia M, Carolei A, Barba C, Frontoni M, Rossi A, Meco G, et al. A placebo-controlled study of the antidepressant activity of moclobemide, a new MAO-A inhibitor. *Pharmacopsychiatry* 1984;**17**(4):122–5.

Davidson 1981 *{published data only}*
Davidson JR, McLeod MN, Turnbull CD, Miller RD. A comparison of phenelzine and imipramine in depressed inpatients. *Journal of Clinical Psychiatry* 1981;**42**(10): 395–7.

DeBattista 2006 *{published data only}*
DeBattista C, Belanoff J, Glass S, Khan A, Horne RL, Blasey C, et al. Mifepristone versus placebo in the treatment of psychosis in patients with psychotic major depression. *Biological Psychiatry* 2006;**15**(12):1343–9.

Ebert 1997 *{published data only}*
Ebert D. Lithium-TCA combination treatment of psychotic depression: comparison with TCA-neuroleptic treatment. *Journal of Clinical Psychopharmacology* 1997;**17**(2):129–30.

Flores 2006 *{published data only}*
Flores BH, Kenna H, Keller J, Solvason HB, Schatzberg AF. Clinical and biological effects of mifepristone treatment for psychotic depression. *Neuropsychopharmacology* 2006;**31**(3): 628–36.

Friedman 1966 *{published data only}*
Friedman AS, Granick S, Cohen HW, Cowitz B. Imipramine (Tofranil vs. placebo in hospitalized psychotic depressives. (a comparison of patients' self-ratings, psychiatrists' ratings and psychological test scores). *Journal of Psychiatric Research* 1966;**4**(1):13–36.

Künzel 2008 *{published data only}*
Künzel HE, Ackl N, Hatzinger M, Held K, Holsboer-Trachsler E, Ising M, et al. Outcome in delusional depression comparing trimipramine monotherapy with a combination of amitriptyline and haloperidol - a double-blind multicenter trial. *Journal of Psychiatric Research* 2009; **43**(7):702–10.

Malison 1999 *{published data only}*
Malison RT, Anand A, Pelton GH, Kirwin P, Carpenter L, McDougle CJ, et al. Limited efficacy of ketoconazole in treatment-refractory major depression. *Journal of Clinical Psychopharmacology* 1999;**19**(5):466–70.

McLaughlin 1969 *{published data only}*
McLaughlin BE. Therapy for psychotic and psychoneurotic depression with thiothixene and perphenazine-amitriptyline (a double-blind clinical study). *Diseases of the Nervous System* 1969;**30**(2):85–8.

Müller 1998 *{published data only}*
Muller-Siecheneder F, Muller MJ, Hillert A, Szegedi A, Wetzel H, Benkert O. Risperidone versus haloperidol and amitriptyline in the treatment of patients with a combined psychotic and depressive syndrome. *Journal of Clinical Psychopharmacology* 1998;**18**(2):111–20.

Navarro 2001 *{published data only}*
Navarro V, Gasto C, Torres X, Marcos T, Pintor L. Citalopram versus nortriptyline in late-life depression: a 12-week randomized single-blind study. *Acta Psychiatrica Scandinavica* 2001;**103**(6):435–40.

Nelson 1984 *{published data only}*
Nelson WH, Kahn A, Orr WW Jr. Phenomenology, neuroendocrine function, and tricyclic antidepressant response. *Journal of Affective Disorders* 1984;**6**(3-4): 297–306.

Spiker 1982 *{published data only}*
Spiker DG, Hanin I, Perel JM, Cofsky JE, Rossi AJ, Sorisio D. Pharmacological treatment of delusional depressives. *Psychopharmacology Bulletin* 1982;**18**(4):184–6.

Zanardi 1998 *{published data only}*
Zanardi R, Franchini L, Gasperini M, Lucca A, Smeraldi E, Perez J. Faster onset of action of fluvoxamine in combination with pindolol in the treatment of delusional depression: a controlled study. *Journal of Clinical Psychopharmacology.* 1998;**18**(6):441–6.

Additional references

Alderson 2004
Alderson P, Green S, Higgins JPT. *Cochrane Reviewer's Handbook 4.2.2*. Chichester: John Wiley & Sons, Ltd, 2004.

APA 1980
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-III), 3rd edition*. Washington, DC: American Psychiatric Association, 1980.

APA 1987
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R), 3rd edition*. Washington, DC: American Psychiatric Association, 1987.

APA 1994
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), 4th edition*. Washington, DC: American Psychiatric Association, 1994.

APA 2000
American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR)*. Washington, DC: American Psychiatric Association, 2000.

APA 2010
American Psychiatric Association. *Practice Guideline for the Treatment of Patients With Major Depressive Disorder (revision)*. Washington, DC: American Psychiatric Association, 2010.

Coryell 1984
Coryell W, Pfohl B, Zimmerman M. The clinical and neuroendocrine features of psychotic depression. *Journal of Nervous and Mental Disease* 1984;**172**(9):521–8.

Coryell 1998
Coryell W. The treatment of psychotic depression. *Journal of Clinical Psychiatry* 1998;**59(Suppl 1)**:22–9.

DerSimonian 1986
DerSimonian R, Laird N. Meta-analysis in clinical trials. *Controlled Clinical Trials* 1986;**7**(3):177–88.

Dutch Guideline 2009
Trimbos, CBO, NHG, LEVV, NVvP, NIP. Dutch Multidisciplinairy Guideline for Depression [Multidisciplinaire Richtlijn Depressie]. Utrecht, The Netherlands, 2009.

Farahani 2012
Farahani A, Corell CU. Are antipsychotics or antidepressants needed for psychotic depression? A systematic review and meta-analysis of trials comparing antidepressant or antipsychotic monotherapy with combination treatment. *Journal of Clinical Psychiatry* 2012;**73**:486–96.

Guy 1976
Guy W. Clinical Global Impression. ECDEU Assessment Manual for Psychopharmacology, revised. National Institute of Mental Health, Rockville, MD 1976.

Hamilton 1960
Hamilton M. A rating scale for depression. *Journal of Neurological and Neurosurgical Psychiatry* 1960;**23**:56–62.

Higgins 2011
Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011.. Available from www.cochrane-handbook.org.

Johnson 1991
Johnson J, Horwath E, Weisman MM. The validity of major depression with psychotic features based on a community study. *Archives of General Psychiatry* 1991;**48**(12):1075–81.

Licht 2008
Licht RW, Gijsman H, Nolen WA, Angst J. Are antidepressants safe in the treatment of bipolar depression? A critical evaluation of their potential risk to induce switch into mania or cycle acceleration. *Acta Psychiatrica Scandinavica* 2008;**118**(5):337–46.

Mahli 2009
Mahli GS, Adams D, Porter R, Wignall A, Lampe L, O'Connor N, et al. Clinical practice recommendations for depression. *Acta Psychiatrica Scandinavica* 2009;**119**:8–26.

Montgomery 1979
Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *British Journal of Psychiatry* 1979;**134**:382–9.

Nelson 1997
Nelson EB, McElroy SL. Psychotic depression a guide to drug choice. *CNS Drugs* 1997;**8**(6):457–73.

NICE 2009
National Institute for Health and Clinical Excellence. *The treatment and management of depression in adults (updated version). National Clinical Practice Guideline 90*. London: HMSO, 2009.

Ohayon 2002
Ohayon MM, Schatzberg AF. Prevalence of depressive episodes with psychotic features in the general population. *American Journal of Psychiatry* 2002;**159**(11):1855–61.

Parker 1992
Parker G, Roy K, Hadzi-Pavlovic D, Pedic F. Psychotic (delusional) depression: a meta-analysis of physical treatments. *Journal of Affective Disorders* 1992;**24**(1):17–24.

Prudic 1990
Prudic J, Sackeim HA, Devanand DP. Medication resistance and clinical response to electroconvulsive therapy. *Psychiatry Research* 1990;**31**(3):287–96.

Prudic 1996
Prudic J, Haskett RF, Mulsant B, Malone KM, Pettinati HM, Stephens S, et al. Resistance to antidepressant medications and short-term clinical response to ECT. *American Journal of Psychiatry* 1996;**153**(8):985–92.

RevMan 2011 [Computer program]
The Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager (RevMan). Version 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011.

Sackeim 2001
Sackeim HA. The definition and meaning of treatment-resistant depression. *Journal of Clinical Psychiatry* 2001;**62 (Suppl 16)**:10–7.

Sadock 2009
Sadock J, Sadock VA, Ruiz P (editors). *Comrehensive Textbook of Psychiatry*. Philadelphia: Lippincott Williams & Wilkins, 2009.

van den Broek 2004a
van den Broek WW, Lely de A, Mulder PGH, Birkenhäger TK, Bruijn J. Effect of antidepressant medication resistance on short-term response to ECT. *Journal of Clinical Psychopharmacology* 2004;**24**(4):400–3.

Wheeler 2000
Wheeler Vega JA, Mortimer AM, Tyson PJ. Somatic treatment of psychotic depression: review and recommendations for practice. *Journal of Clinical Psychopharmacology* 2000;**20**(5):504–19.

Wijkstra 2005
Wijkstra J, Lijmer J, Balk F, Geddes J, Nolen WA. Pharmacological treatment for psychotic depression. *Cochrane Database of Systematic Reviews* 2005. DOI: 10.1002/14651858.CD004044.pub2

Wijkstra 2007
Wijkstra J, Schubart CD, Nolen WA. Treatment of unipolar psychotic depression: the use of evidence in practice guideline. *World Journal of Biological Psychiatry* 2007;**24**: 1–7.

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Anton 1990

Methods	Randomised, double-blind comparison
Participants	No explicit use of structured interview *DSM-III* criteria; psychotic depressive episode HRSD 17 > 18 Inpatients. No data about prior treatment of current episode
Interventions	Amoxapine versus amitriptyline + perphenazine. 300 to 500 mg versus 150 to 250 mg + 24 to 40 mg No blood levels Five days placebo period. Additional medication in these five days: lorazepam or oxazepam in 'low dose' Treatment period: four weeks. Additional medication is not mentioned in these four weeks
Outcomes	Dichotomous data: author defined: Response is reduction of HRSD-17 > 50%. No remission data Continuous data: symptom reduction: no ITT data; global response: no ITT data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: yes (two in ami + per) Mortality rate: 0
Notes	56 participants provided informed consent. 10 dropped out in washout before receiving active medication (four refused and six improved substantially); 46 participants were randomly assigned 46 participants: four dropouts in both groups (total eight). Unclear how many bipolar participants among these eight dropouts; 38 participants were analysed, including six bipolar participants 6/38 bipolar = 15.8% ITT responders: amoxapine 12/21 and ami + per 17/25 (instead of 12/17 and 17/21) Dropouts after random assignment: 9/21 and 7/25 Author had no additional data available See also 1993 J Aff Disorders 28:125-131 (same data set)

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	As reported: 'Patients were randomly assigned in a double blind fashion'
Allocation concealment (selection bias)	Unclear risk	No information

Anton 1990 (*Continued*)

Blinding (performance bias and detection bias) of participants	Low risk	As reported: 'Double blind treatment with identical capsules'
Blinding (performance bias and detection bias) of personnel	Low risk	As reported: 'Double blind treatment with identical capsules'
Blinding (performance bias and detection bias) of outcome assessors	Unclear risk	Probably yes. No explicit data
Incomplete outcome data (attrition bias) All outcomes	High risk	46 participants were randomly assigned. In the publication, only those participants who completed at least two weeks of active medication were analysed. Four dropouts in both groups (total eight)
Selective reporting (reporting bias)	Unclear risk	No protocol available. Generally accepted outcomes have been used
Other bias	High risk	Unclear how many bipolar participants were present among these eight dropouts; 38 participants were analysed, including six bipolar participants. 6/38 bipolar = 15.8%. No additional data available to exclude bipolar participants from reanalysis We reanalyzed the data with ITT responders (intention-to-treat; dropouts included): amoxapine 12/21 and ami + per 17/25 (instead of 12/17 and 17/21). ITT dropouts after random assignment: 9/21 and 7/25

Bruijn 1996

Methods	Randomised, double-blind comparison
Participants	Use of checklist with *DSM-III-R* criteria SADS depression portion was performed in the presence of a second psychiatrist *DSM-III-R* depressive episode; excluded psychotic depression with hallucinations HRSD-17 > 17 Inpatients. Subgroup psychotic depression. Probably only with delusions 51% of included participants were adequately pretreated during the current episode: adequate dose of an antidepressant during at least four weeks
Interventions	Imipramine versus mirtazapine; 37.5 to 450 mg imipramine (blood level: 199 to 400 ng/mL) versus 40 to 100 mg mirtazapine (blood level 49 to 93 ng/mL) Washout: three days medication free and four days placebo

Bruijn 1996 (*Continued*)

	Additional medication: one to six tablets a day containing 45 mg of an extract of valerian, lorazepam 1 to 5 mg a day or haloperidol 1 to 15 mg a day Treatment period: four weeks after predefined blood levels reached (mirtazapine group: five to 21 days; imipramine group: seven to 25 days)
Outcomes	Dichotomous data: author defined: Response is reduction of HRSD-17 ≥ 50%. No remission data Continuous data: symptom reduction: no ITT data; global response: no ITT data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: no ITT data in subgroup Mortality rate: 0
Notes	Worse responding in a group leads to more participants given haloperidol 107 participants included; six bipolar; 10 dropouts Subgroup: MDD psychotic; 30 (15 mirtazapine and 15 imipramine) Mirtazapine group: seven haloperidol treatment (six non-responders, one responder) Imipramine group: two haloperidol treatment (two non-responders) Participants treated with haloperidol by us counted as dropouts Mirtazapine group: one dropout + seven haloperidol treatment = 8/18; imipramine group: two dropouts + two haloperidol treatment = 4/15 Additional information from author included

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	As reported: 'Patients were randomly allocated to a double blind treatment'
Allocation concealment (selection bias)	Unclear risk	No explicit information
Blinding (performance bias and detection bias) of participants	Low risk	As reported: 'Identical capsules. Dose adjustment by an independent psychiatrist on the basis of blood levels'
Blinding (performance bias and detection bias) of personnel	Low risk	As reported: 'Identical capsules. Dose adjustment by an independent psychiatrist'
Blinding (performance bias and detection bias) of outcome assessors	Low risk	Side effects were not systematically rated to prevent bias towards unblinding. After completion of the study, the research psychiatrist guessed the medication: 46 correct and 37 incorrect
Incomplete outcome data (attrition bias) All outcomes	Low risk	None

Bruijn 1996 (*Continued*)

Selective reporting (reporting bias)	Unclear risk	No protocol available. Generally accepted outcomes have been used
Other bias	High risk	Participants with psychotic depression with hallucinations were excluded. So only participants with psychotic depression with delusions were included in the reanalyzed subgroup We reanalyzed the data in the subgroup with psychotic depression. We counted as dropout: one participant with bipolar disorder, nine participants with haloperidol treatment (seven in mirtazapine group and two in imipramine group) Worse responding in psychotic depression leads in this study to more open co-treatment with haloperidol 1 to 15 mg, especially in the mirtazapine group. Only one of these nine participants (mirtazapine group) was a responder. So haloperidol probably was not instrumental in the recovery of those participants

Meyers 2009

Methods	Randomised double-blind study
Participants	259 participants; *DSM-IV-TR* psychotic depression; 18 years of age or older; HAM-D $\geq$ 21 and SADS delusional severity rating $\geq$ 3 Inpatients
Interventions	12 weeks treatment with olanzapine + placebo and olanzapine + sertraline
Outcomes	Remission rates (HAM-D 17 $\leq$ 10 and SADS delusional item score = 1)
Notes	53% dropout in olanzapine arm and 37% dropout in olanzapine + sertraline arm

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	As reported: 'Computer generated randomisation list'
Allocation concealment (selection bias)	Unclear risk	No further data

Meyers 2009 (*Continued*)

Blinding (performance bias and detection bias) of participants	Low risk	Study was double blind (reported as: 'sertraline and placebo under double-blind conditions')
Blinding (performance bias and detection bias) of personnel	Low risk	Well described double blinding. 'Sertraline and placebo under double-blind conditions'. As reported: 'Investigators and raters were blind to treatment assignment'
Blinding (performance bias and detection bias) of outcome assessors	Low risk	Well described double blinding. 'Sertraline and placebo under double-blind conditions'. As reported: 'Investigators and raters were blind to treatment assignment'
Incomplete outcome data (attrition bias) All outcomes	Low risk	None
Selective reporting (reporting bias)	Low risk	Protocol available. Generally accepted outcomes have been used
Other bias	Unclear risk	Relatively high dropout rate with significant difference between treatment groups.(53% olanzapine and 37% olanzapine + sertraline) Patients with only hallucinations excluded

Mulsant 2001

Methods	Randomised, double-blind comparison
Participants	Clinical interview, Brief Psychiatric Rating Scale, Global Assessment Scale and a consensus conference were used for diagnosis *DSM-III-R*; psychotic major depressive episode (manic episode in history excluded). HAM-D-17 > 17 Age > 50 years Inpatients No data about prior treatment of current episode
Interventions	Nortriptyline versus nortriptyline + perphenazine Open nortriptyline until therapeutic plasma level (target 100 ng/mL); once between 50 and 150 ng/mL, random assignment followed Mean doses: nortriptyline 76 mg versus nortriptyline 63 mg + perphenazine 19 mg Mean blood levels: 101 ng/mL versus 120 + 4 ng/mL Additional medication: lorazepam as needed Treatment period: after random assignment 2 to 16 weeks (total treatment at least four weeks) "After a washout of other psychotropic medication except lorazepam ...." It is unclear how long this washout has been

Mulsant 2001 (*Continued*)

Outcomes	Dichotomous data: author defined: Response is HAMD-17 < 11 and BPRS (11, 12, 15) 1 or 2. No remission data Continuous data: symptom reduction: no ITT data; global response: no ITT data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: nortriptyline + perphenazine 1/17; nortriptyline + placebo 2/19 Mortality rate: 0
Notes	54 participants included; 16 dropouts: three improved on nortriptyline, three adverse effects and nine administrative reasons; 36 participants randomly assigned. This is by procedure a selected group: responders on nortriptyline and participants with adverse effects and with other reasons are excluded (28%) Open nortriptyline (eight to 21 days; median two weeks); once between 50 and 150 ng/mL, random assignment followed Responder somewhere between two and 16 weeks after randomisation (median nine weeks); three dropouts in both groups after random assignment. These are excluded by the author and included by us for ITT analysis

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	No information. As reported: 'Patient[s] were randomly allocated to a double blind treatment'
Allocation concealment (selection bias)	Unclear risk	No information. As reported: 'Patient[s] were randomly allocated to a double blind treatment'
Blinding (performance bias and detection bias) of participants	Low risk	As reported: 'Double blind treatment. No further data'
Blinding (performance bias and detection bias) of personnel	Low risk	'Dose adjustments by non blinded psychiatrist who were not involved in the care'
Blinding (performance bias and detection bias) of outcome assessors	Low risk	As reported: 'Double blind treatment'
Incomplete outcome data (attrition bias) All outcomes	Low risk	None
Selective reporting (reporting bias)	Unclear risk	No protocol available. Generally accepted outcomes has been used

Mulsant 2001 (*Continued*)

Other bias	High risk	No outcome data about prescription of 'lorazepam as needed' EPS rating could have led to blinding bias Participants treated with only nortriptyline (+ placebo) were excluded after four weeks without improvement, and participants treated with nortriptyline + perphenazine after four + two weeks

Rothschild 2004a

Methods	Randomised, double-blind comparison Random assignment was 2:2:1 for olanzapine, placebo and olanzapine fluoxetine combination, respectively
Participants	124 participants; *DSM-IV* diagnosis (unclear how) Major depression with psychotic features Inpatients for at least one week. No data about prior treatment of current episode
Interventions	2004a: olanzapine (5 to 20 mg, clinically titrated; mean 11.9 mg) versus olanzapine (5 to 20 mg, clinically titrated; mean 12.4 mg) plus fluoxetine (20 to 80 mg, clinically titrated; mean 23.5 mg) versus placebo 2004b: same procedure: olanzapine (mean 14.0 mg) versus olanzapine (mean 13.9 mg) plus fluoxetine (mean 22.6 mg) Three to nine days screening; probably no washout period Treatment period: eight weeks Additional medication: 30 mg a day diazepam equivalent for no more than five consecutive days or 10 cumulative days
Outcomes	Dichotomous data: author defined: Response is reduction of HAMD-24 $\geq$ 50% at endpoint. Remission is HAMD-24 $\leq$ 8 for two consecutive visits Continuous data: symptom reduction: no ITT data; global response: no ITT data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: no ITT data Mortality rate: probably 0
Notes	Washout unclear 23 investigators randomly assigned at least one participant. Excluded patient characteristics not described Dropouts in study a: 28%; lost before baseline + one visit: 7% (were excluded from results; included in our data); 24% in study a are LOCF (last observation carried forward; in our data not counted as dropouts); some of these LOCFs are counted as responders; total non-completers (LOCF included) 28 + 7 + 24 = 59% Dropouts in study b: 38%; lost before baseline + one visit: 9% (were excluded from results; included in our data); LOCF in study b: 6%; total non-completers 28 + 9 + 6 = 53%

	Completers in study a: 41%; in study b: 47%	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	As reported: 'Patients were randomly allocated'; no further information
Allocation concealment (selection bias)	Unclear risk	As reported: 'Patients were randomly allocated'; no further information
Blinding (performance bias and detection bias) of participants	Low risk	As reported: 'Double blind therapy. Dose adjustments in all study arms with 'capsules' (assuming identical capsules because the study is double blind)'
Blinding (performance bias and detection bias) of personnel	Low risk	As reported: 'Double blind therapy'
Blinding (performance bias and detection bias) of outcome assessors	Unclear risk	No explicit information
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Dropouts described in general terms. Very high dropout rate
Selective reporting (reporting bias)	Unclear risk	No protocol available. According to the authors, the olanzapine-fluoxetine group was designed as a exploratory pilot arm. However, in the conclusions, it is stated that an olanzapine/fluoxetine combination was a well-tolerated treatment associated with significant and quick reduction in depressive (and psychotic) symptoms in one trial. With ITT data, this difference is seen in one study to be not statistically significant, and in the other study barely significant. Pooling of these two studies would result in no significance The authors discuss as a limitation the absence of a fluoxetine arm. They state that they cannot rule out that the effect of fluoxetine/olanzapine was due to fluoxetine. So this should have been mentioned in the conclusion

Rothschild 2004a (*Continued*)

Other bias	High risk	High dropout rate of 34.7% reduces the internal validity of the study High placebo response is contradictory to the literature

Rothschild 2004b

Methods	Randomised, double-blind comparison Random assignment was 2:2:1 for olanzapine, placebo and olanzapine fluoxetine combination, respectively
Participants	124 participants. *DSM-IV* diagnosis (unclear how) Major depression with psychotic features Inpatients for at least one week. No data about prior treatment of current episode
Interventions	2004a: olanzapine (5 to 20 mg, clinically titrated; mean 11.9 mg) versus olanzapine (5 to 20 mg, clinically titrated; mean 12.4 mg) plus fluoxetine (20 to 80 mg, clinically titrated; mean 23.5 mg) versus placebo 2004b: same procedure: olanzapine (mean 14.0 mg) versus olanzapine (mean 13.9 mg) plus fluoxetine (mean 22.6 mg) Three to nine days screening; probably no washout period Treatment period: eight weeks Additional medication: 30 mg a day diazepam equivalent for no more than five consecutive days or 10 cumulative days
Outcomes	Dichotomous data: author defined: Response is reduction of HAMD-24 $\geq$ 50% at endpoint. Remission is HAMD-24 $\leq$ 8 for two consecutive visits Continuous data: symptom reduction: no ITT data; global response: no ITT data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: no ITT data Mortality rate: probably 0
Notes	Washout unclear 23 investigators randomly assigned at least one participant. Excluded patient characteristics not described Dropouts in study a: 28%; lost before baseline + one visit: 7% (were excluded from results; included in our data); 24% in study a are LOCF (last observation carried forward; in our data not counted as dropouts). Some of these LOCFs are counted as responders; total non-completers (LOCF included) 28 + 7 + 24 = 59% Dropouts in study b: 38%; lost before baseline + one visit: 9% (were excluded from results; included in our data); LOCF in study b: 6%; total non-completers 28 + 9 + 6 = 53% Completers in study a: 41%; in study b: 47%

Risk of bias

Rothschild 2004b (*Continued*)

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	As reported: 'Patients were randomly allocated'; no further information
Allocation concealment (selection bias)	Unclear risk	As reported: 'Patients were randomly allocated'; no further information
Blinding (performance bias and detection bias) of participants	Low risk	As reported: 'Double blind therapy. Dose adjustments in all study arms with 'capsules' (assuming identical capsules because the study is double blind)'
Blinding (performance bias and detection bias) of personnel	Low risk	As reported: 'Double blind therapy'
Blinding (performance bias and detection bias) of outcome assessors	Unclear risk	No explicit information
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Dropouts described in general terms. Very high dropout rate
Selective reporting (reporting bias)	Unclear risk	No protocol available. According to the authors, the olanzapine-fluoxetine group was designed as a exploratory pilot arm. However, in the conclusions, it is stated that an olanzapine/fluoxetine combination was a well-tolerated treatment associated with significant and quick reduction in depressive (and psychotic) symptoms in one trial. With ITT data, this difference is seen in one study to be not statistically significant, and in the other study to be barely significant. Pooling of these two studies would result in no significance. The authors discuss as a limitation the absence of a fluoxetine arm. They state that they cannot rule out that the effect of fluoxetine/olanzapine was due to fluoxetine. So this should have been mentioned in the conclusion
Other bias	High risk	High dropout rate of 47.2% reduces the internal validity of the study High placebo response is contradictory to the literature

Spiker 1985

Methods	Randomised, double-blind comparison Random assignment procedure described in part Blinding adequately described
Participants	SADS and RDC criteria for major depressive disorder, primary subtype and psychotic subtype (only with delusions); bipolar participants included Severity rating 4 or greater on 6-point scale in the SADS that rates severity of delusion HRSD-17 > 14 Inpatients No data about prior treatment of current episode
Interventions	Three groups: perphenazine versus amitriptyline versus amitriptyline + perphenazine Doses: perphenazine mean 50 mg versus amitriptyline mean 218 mg versus amitriptyline mean 170 mg + perphenazine mean 54 mg Blood levels: perphenazine 19 to 113 ng/mL versus amitriptyline (+ nortriptyline) 130 to 500 ng/mL versus amitriptyline 157 to 690 ng/mL + perphenazine 18 to 128 ng/mL Seven days drug free Treatment period: four weeks Additional medication: benztropine mesylate 4 mg
Outcomes	Dichotomous data: author defined: Response is HRSD-17 < 7 and delusional rating score = 1 (6-point scale in the SADS). No remission data (definition of response is definition of remission) Continuous data: symptom reduction: no ITT data; global response: no ITT data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: amitriptyline + perphenazine 2/22, perphenazine 1/17 Mortality rate: 0
Notes	Only participants with delusions Seven drop out in ITT (in the original data, dropouts are excluded from the analysis); response data ITT 3/17 (original 3/16); 7/19 (7/17); 14/22 (14/18) 9/58 = 15.5% bipolar participants in analysis. Because of lack of data, we were not able to exclude these bipolar participants from the analysis

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	As reported: 'The hospital pharmacist assigned patients randomly'
Allocation concealment (selection bias)	Low risk	Probably yes: 'The hospital pharmacist assigned patients randomly'
Blinding (performance bias and detection bias) of participants	Low risk	As reported: 'The hospital pharmacist assigned patients randomly. All tablets looked identical' 'All raters and floor staff and the patient

Spiker 1985 (*Continued*)

		were blind to the patient's drug treatment and the plasma-level data'
Blinding (performance bias and detection bias) of personnel	Low risk	As reported: 'The hospital pharmacist assigned patients randomly. All tablets looked identical' 'All raters and floor staff and the patient were blind to the patient's drug treatment and the plasma-level data'
Blinding (performance bias and detection bias) of outcome assessors	Low risk	As reported: 'The hospital pharmacist assigned patients randomly. All tablets looked identical' 'All raters and floor staff and the patient were blind to the patient's drug treatment and the plasma-level data'
Incomplete outcome data (attrition bias) All outcomes	Low risk	None
Selective reporting (reporting bias)	Unclear risk	No protocol available. Generally accepted outcomes have been used
Other bias	Unclear risk	Only participants with delusions are included 9/58 = 15.5% bipolar participants in analysis. Because of lack of data, we were not able to exclude these bipolar participants from the analysis We reanalyzed the data to ITT

Spiker 1988

Methods	Reanalyzing two studies (not including the data of the Spiker 1985 study) Randomised, double-blind, placebo-controlled Randomisation procedure not explicitly described Blinding adequately described in original studies
Participants	Re-diagnosing by using *DSM-III* criteria Major depressive disorder *DSM-III* HRSD-17 > 14 (30 or more based on the sum of two raters) Inpatients. Subgroup psychotic participants No data about prior treatment of current episode
Interventions	Amitriptyline versus placebo Three days 50; four days 100; seven days 150, 14 days 200 mg amitriptyline (at least three weeks ≥ 150 mg) Blood levels: unknown

Spiker 1988 (*Continued*)

	Extra medication: none Two weeks drug-free washout period One week placebo (single-blind); total period of three weeks drug free Treatment period: four weeks
Outcomes	Dichotomous data: author defined: Response is HRSD-17 < 7 (< 14/2) + not psychotic or HRSD-17 = 6.5 to 9.5 (13/2 to 19 /2) + not psychotic + 1/3 or less of entering score. Remission data not specified Continuous data: symptom reduction: no data; global response: no data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: no data Mortality rate: 0 (no data)
Notes	20% response in two-week drug-free period (psychotic + non-psychotic); no data about psychotic versus non-psychotic in these two weeks Four weeks treatment; only two weeks 200 mg; no blood levels Subgroup of 27 participants with psychotic depression. Amitriptyline 14; placebo 13 Dropouts four (amitriptyline) and three (placebo) are excluded from analysis by the authors. Responders amitriptyline 4/10 and placebo 0/10. ITT responders: amitriptyline 4/14 and placebo 0/13

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	As reported: 'Patients were randomly assigned'
Allocation concealment (selection bias)	Unclear risk	No data
Blinding (performance bias and detection bias) of participants	Low risk	As reported: 'All patients received 4 identical capsules daily: patients and staff were blind'
Blinding (performance bias and detection bias) of personnel	Low risk	As reported: 'All patients received 4 identical capsules daily: patients and staff were blind'
Blinding (performance bias and detection bias) of outcome assessors	Unclear risk	No data
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	None
Selective reporting (reporting bias)	Unclear risk	No protocol available. Generally accepted outcomes have been used

Spiker 1988 (*Continued*)

Other bias	High risk	Participants were retrospectively re-diagnosed. Psychotic depression and non-psychotic depression were included and randomly assigned. We used the data about psychotic participants 14 days drug-free period (20% remission with no further data) + one week placebo before random assignment could be due to low placebo response

van den Broek 2004

Methods	Randomised, double-blind comparison
Participants	*DSM-IV* diagnosis major depressive disorder; assessed with the depression part of the SADS HRSD-17 > 16 Inpatients; subgroup of psychotic depression 39% of all included participants were pretreated with an SSRI and 22.7% with a TCA, but none as inpatients with adequate plasma level for at least four weeks during the present episode
Interventions	Imipramine versus fluvoxamine Four days placebo washout Predefined blood levels. Imipramine 150 to 450 mg (blood level imipramine + desimipramine 192 to 521). Fluvoxamine 150 to 1800 mg (blood level 109 to 325 ng/mL). Treatment period: four weeks after reaching predefined blood levels. Additional medication: one to six tablets a day containing 45 mg of an extract of valerian, lorazepam 1 to 3 mg a day or haloperidol 1 to 10 mg a day
Outcomes	Dichotomous data: author defined: Response is reduction of HRSD-17 $\geq$ 50%. Remission is HRSD-17 < 8
Notes	Subgroup with psychotic features. Some participants in this subgroup had been treated with haloperidol (by us counted as dropouts). Worse responding in a group leads to more participants who were given haloperidol We used additional psychotic subgroup data from the author

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	As reported: 'A computer generated randomisation list was used'
Allocation concealment (selection bias)	Unclear risk	No specific data

van den Broek 2004 (*Continued*)

Blinding (performance bias and detection bias) of participants	Low risk	As reported: 'Tablets identical in appearance, weight and taste were administered. Preparation of the tablets was done by the pharmacist. The treating physician received blood level data in percentages'
Blinding (performance bias and detection bias) of personnel	Low risk	As reported: 'Tablets identical in appearance, weight and taste were administered. Preparation of the tablets was done by the pharmacist. The treating physician received blood level data in percentage'
Blinding (performance bias and detection bias) of outcome assessors	Low risk	As reported: 'The treating physicians were not involved in the ratings'
Incomplete outcome data (attrition bias) All outcomes	Low risk	None
Selective reporting (reporting bias)	Unclear risk	No protocol available. Generally accepted outcomes have been used
Other bias	High risk	Worse responding in a group leads to more participants who were given haloperidol We reanalyzed the data: Participants treated with haloperidol were counted as dropouts

Wijkstra 2010

Methods	Randomised double-blind study	
Participants	*DSM-IV*-defined psychotic depression Inpatients	
Interventions	Seven weeks of treatment with imipramine (plasma levels 200 to 300 μg/L), venlafaxine (375 mg/d), venlafaxine + quetiapine (375 mg/d + 600 mg/d)	
Outcomes	Dichotomous data: author defined (response). Response is $\geq$ 50% decrease in HAM-D 17 scores from baseline to study endpoint, and a final HAM-D score $\leq$ 14. Remission is HAMD $\leq$ 7 (not predefined)	
Notes	No quetiapine arm. Inclusion did not reach planned number (122 i.s.o. 155) Relatively low dropout rate (22/122 = 18%)	
Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**

Wijkstra 2010 (*Continued*)

Random sequence generation (selection bias)	Low risk	As reported: 'Randomisation was executed centrally using a computer-generated randomisation list: randomly permuted blocks of size six'
Allocation concealment (selection bias)	Low risk	As reported: 'Randomisation was executed *centrally*'
Blinding (performance bias and detection bias) of participants	Low risk	The study was double-blind. Blood was collected from each participant (only imipramine blood level was assessed). Treatment guesses were analysed and indicated high preservation of blindness
Blinding (performance bias and detection bias) of personnel	Low risk	Double-blind study. Blood was collected from each participant (only imipramine blood level was assessed). Treatment guesses were analysed and indicated high preservation of blindness
Blinding (performance bias and detection bias) of outcome assessors	Low risk	As reported: 'Blindness was checked and high. All dose adjustments were done centrally by an independent psychiatrist'
Incomplete outcome data (attrition bias) All outcomes	Low risk	None
Selective reporting (reporting bias)	Low risk	Protocol available. Generally accepted outcomes have been used
Other bias	Unclear risk	122 participants included i.s.o. with the planned 155 resulting in loss of power Post hoc remission as secondary outcome measure

Zanardi 1996

Methods	Randomised, double-blind comparison
Participants	The SCID patient version was used for some participants but not for all (reply of author to letter to editor) *DSM-III-R* criteria; psychotic depressive episode No HRSD criteria described at inclusion Inpatients No data about prior treatment of current episode
Interventions	Sertraline versus paroxetine Dose: 150 mg versus 50 mg from day 8 Blood levels: unknown Additional medication: flurazepam < 30 mg (bipolar participants additional medication lithium; bipolar participants are excluded from our data) One week medication free (single-blind placebo period) Treatment period: five weeks

Outcomes	Dichotomous data: author defined: Response is HRSD-21 < 8 + DDERS (Dimensions of Delusional Experience RS) = 0. Remission data not specified Continuous data: symptom reduction: no data; global response: no data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: same as overall dropout Mortality rate: 0
Notes	5/14 dropouts in paroxetine group and 0/18 in sertraline group Bipolar participants could be excluded from our analysis

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	As reported: 'Patients were randomly assigned to two therapy groups'
Allocation concealment (selection bias)	Unclear risk	No data
Blinding (performance bias and detection bias) of participants	Unclear risk	As reported: 'Patients were randomly assigned'. In title and abstract: 'Double-blind controlled trial. No information about methods of blinding'
Blinding (performance bias and detection bias) of personnel	Unclear risk	As reported: 'Double-blind controlled trial'
Blinding (performance bias and detection bias) of outcome assessors	Unclear risk	No data
Incomplete outcome data (attrition bias) All outcomes	Low risk	None
Selective reporting (reporting bias)	Unclear risk	No protocol available. Generally accepted outcomes have been used
Other bias	Unclear risk	We reanalyzed the data by excluding bipolar participants Difference in dropout is high: 5/14 versus 0/18

Zanardi 2000

Methods	Randomised, double-blind comparison
Participants	Unclear diagnosing procedure *DSM-IV* criteria; psychotic depressive episode No HRSD criteria described at inclusion Inpatients No data about prior treatment of current episode
Interventions	Venlafaxine versus fluvoxamine Dose: 300 mg versus 300 mg from day 8 Blood levels: unknown Additional medication: flurazepam < 30 mg One week medication free (single-blind placebo period) Treatment period: five weeks
Outcomes	Dichotomous data: author defined: Response is HRSD-21 < 9 + DDERS (Dimensions of Delusional Experience RS) = 0. Remission data not specified Continuous data: symptom reduction: no ITT data; global response: no ITT data; QOL: no data Overall dropout rate: yes Dropout due to adverse effects: same as overall dropout rate Mortality rate: 0
Notes	We used additional data from the author to exclude the bipolar participants from the analysis Included 22 participants with major depressive disorder (MDD) with psychotic features. Responders in venlafaxine group 6/11 MDD. Responders fluvoxamine group 9/11 MDD. No dropouts in fluvoxamine group. Two dropouts in venlafaxine group

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	As reported: 'Randomization was performed by a computer-generated schedule'
Allocation concealment (selection bias)	Unclear risk	Randomisation method not described Blinding not explicitly described. No data
Blinding (performance bias and detection bias) of participants	Unclear risk	As reported: 'Patients were randomly assigned' 'Double-blind controlled study'
Blinding (performance bias and detection bias) of personnel	Unclear risk	'Double-blind controlled study', but unclear whether double-blind includes personnel

Zanardi 2000 (*Continued*)

Blinding (performance bias and detection bias) of outcome assessors	Low risk	As reported: 'Raters were blind to treatment'
Incomplete outcome data (attrition bias) All outcomes	Low risk	None
Selective reporting (reporting bias)	Unclear risk	No protocol available. Generally accepted outcomes have been used
Other bias	Unclear risk	We reanalyzed the data by excluding bipolar participants with additional data from the author

DSM: Diagnostic and Statistical Manual of Mental Disorders.
HRSD/HAM-D: Hamilton Rating Scale for Depression.
ITT: Intention-to-Treat.
LOCF: Last Observation Carried Forward.
MDD: Major Depressive Disorder.
QOL: Quality of Life.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Belanoff 2001	Only four days of treatment
Bellini 1994	25% bipolar participants (in each group three bipolar participants) The author did not respond to our request for additional information
Blasey 2009	Impossible to compare two defined pharmacological treatments
Blasey 2011	Impossible to compare two defined pharmacological treatments
Cassacchia 1984	'Unipolar psychotic depression' is probably 'manic depressive psychosis, depressive type' *(ICD9)*. This is not the same as 'psychotic depression' Number of bipolar participants is not clear Dropouts not in results. It is not possible to extract ITT response data Reason for exclusion: unclear diagnosis, number of bipolar participants unclear, ITT data not available
Davidson 1981	Reason for exclusion: unclear diagnosis and short treatment period
DeBattista 2006	Impossible to compare two defined pharmacological treatments. 48.3% + 12.9% = 61.2% HAMD response with placebo after one week

(*Continued*)

Ebert 1997	Randomisation not adequate, open study
Flores 2006	Impossible to compare two defined pharmacological treatments and treatment only seven days
Friedman 1966	No comparable diagnostic procedure. No data about MDD subgroup Dropouts have been excluded
Künzel 2008	No ITT data; bipolar participants 17.5% in per-protocol data; continued treatment with lithium, valproic acid
Malison 1999	Only three psychotic participants
McLaughlin 1969	Diagnosis unclear
Müller 1998	In this subgroup, no data are given about responders, bipolar participants and dropouts The author did not respond to our request for additional information
Navarro 2001	Citalopram versus nortriptyline Subgroup with nine psychotic depressive episodes Reason for exclusion: This subgroup was also treated with haloperidol. No data available about this subgroup The author did not respond to our request for additional information
Nelson 1984	Unknown from data in which group the responders are located (imipramine or ami). So comparison is impossible
Spiker 1982	Pre-published data from the 1985 study
Zanardi 1998	30.5 % bipolar participants

DATA AND ANALYSES

Comparison 1. Antidepressant versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical response	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.1 Amitriptyline versus placebo	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Dropouts	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.1 Amitriptyline versus placebo	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 2. Antipsychotic versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical response	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.1 Olanzapine versus placebo	2	201	Risk Ratio (M-H, Fixed, 95% CI)	1.13 [0.74, 1.73]
2 Dropouts	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.1 Olanzapine versus placebo	2	201	Risk Ratio (M-H, Fixed, 95% CI)	0.79 [0.57, 1.08]

Comparison 3. Antidepressant versus antidepressant

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical response	5		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.1 Imipramine versus venlafaxine	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.2 Imipramine versus mirtazapine	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.3 Imipramine versus fluvoxamine	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.4 Fluvoxamine versus venlafaxine	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
1.5 Sertraline versus paroxetine	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Dropouts	5		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.1 Imipramine versus venlafaxine	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

2.2 Imipramine versus mirtazapine	1	Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
2.3 Imipramine versus fluvoxamine	1	Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
2.4 Fluvoxamine versus venlafaxine	1	Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
2.5 Sertraline versus paroxetine	1	Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 4. Antidepressant versus antipsychotic

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical response	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.1 Amitriptyline versus perphenazine	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]
2 Dropouts	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
2.1 Amitriptyline versus perphenazine	1		Risk Ratio (M-H, Fixed, 95% CI)	0.0 [0.0, 0.0]

Comparison 5. Antidepressant plus antipsychotic versus placebo

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical response	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.1 Fluoxetine + olanzapine versus placebo	2	148	Risk Ratio (M-H, Fixed, 95% CI)	1.86 [1.23, 2.82]
2 Dropouts	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
2.1 Fluoxetine + olanzapine versus placebo	2	148	Risk Ratio (M-H, Fixed, 95% CI)	0.75 [0.48, 1.18]

Comparison 6. Antidepressant plus antipsychotic versus placebo plus antipsychotic

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical response	4	447	Risk Ratio (M-H, Fixed, 95% CI)	1.83 [1.40, 2.38]
1.1 amitriptyline + perphenazine versus perphenazine	1	39	Risk Ratio (M-H, Fixed, 95% CI)	3.61 [1.23, 10.56]
1.2 Fluoxetine + olanzapine versus olanzapine	2	149	Risk Ratio (M-H, Fixed, 95% CI)	1.64 [1.10, 2.44]

1.3 Olanzapine + sertraline versus olanzapine	1	259	Risk Ratio (M-H, Fixed, 95% CI)	1.76 [1.21, 2.54]
2 Dropouts	4	447	Risk Ratio (M-H, Fixed, 95% CI)	0.79 [0.63, 1.01]
2.1 Amitriptyline + perphenazine versus perphenazine	1	39	Risk Ratio (M-H, Fixed, 95% CI)	3.09 [0.38, 25.19]
2.2 Fluoxetine + olanzapine versus olanzapine	2	149	Risk Ratio (M-H, Fixed, 95% CI)	0.95 [0.59, 1.53]
2.3 Olanzapine + sertraline versus olanzapine	1	259	Risk Ratio (M-H, Fixed, 95% CI)	0.70 [0.53, 0.92]

Comparison 7. Antidepressant plus antipsychotic versus placebo plus antidepressant

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical response	4	245	Risk Ratio (M-H, Fixed, 95% CI)	1.42 [1.11, 1.80]
1.1 Nortriptyline + perphenazine versus nortriptyline	1	36	Risk Ratio (M-H, Fixed, 95% CI)	1.12 [0.49, 2.53]
1.2 Venlafaxine + quetiapine versus venlafaxine	1	59	Risk Ratio (M-H, Fixed, 95% CI)	1.95 [1.13, 3.37]
1.3 Amitriptyline + perphenazine versus amitriptyline	1	41	Risk Ratio (M-H, Fixed, 95% CI)	1.73 [0.89, 3.37]
1.4 Amitriptyline + perphenazine versus amoxapine	1	46	Risk Ratio (M-H, Fixed, 95% CI)	1.19 [0.75, 1.88]
1.5 Venlafaxine + quetiapine versus imipramine	1	63	Risk Ratio (M-H, Fixed, 95% CI)	1.27 [0.84, 1.93]
2 Dropouts	4	245	Risk Ratio (M-H, Fixed, 95% CI)	0.91 [0.55, 1.50]
2.1 Nortriptyline + perphenazine versus nortriptyline	1	36	Risk Ratio (M-H, Fixed, 95% CI)	1.12 [0.26, 4.81]
2.2 Venlafaxine + quetiapine versus venlafaxine	1	59	Risk Ratio (M-H, Fixed, 95% CI)	0.73 [0.22, 2.46]
2.3 Amitriptyline + perphenazine versus amitriptyline	1	41	Risk Ratio (M-H, Fixed, 95% CI)	1.73 [0.35, 8.41]
2.4 Amitriptyline + perphenazine versus amoxapine	1	46	Risk Ratio (M-H, Fixed, 95% CI)	0.65 [0.29, 1.45]
2.5 Venlafaxine + quetiapine versus imipramine	1	63	Risk Ratio (M-H, Fixed, 95% CI)	1.14 [0.38, 3.47]

Comparison 8. Antidepressant plus antipsychotic versus placebo plus the same antidepressant

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1 Clinical response	3	157	Risk Ratio (M-H, Fixed, 95% CI)	1.70 [1.19, 2.43]
1.1 Nortriptyline + perphenazine versus nortriptyline	1	36	Risk Ratio (M-H, Fixed, 95% CI)	1.12 [0.49, 2.53]
1.2 Venlafaxine + quetiapine versus venlafaxine	1	80	Risk Ratio (M-H, Fixed, 95% CI)	1.98 [1.20, 3.24]
1.3 Amitriptyline + perphenazine versus amitriptyline	1	41	Risk Ratio (M-H, Fixed, 95% CI)	1.73 [0.89, 3.37]
2 Dropouts	3	157	Risk Ratio (M-H, Fixed, 95% CI)	1.04 [0.52, 2.07]
2.1 Nortriptyline + perphenazine versus nortriptyline	1	36	Risk Ratio (M-H, Fixed, 95% CI)	1.12 [0.26, 4.81]
2.2 Venlafaxine + quetiapine versus venlafaxine	1	80	Risk Ratio (M-H, Fixed, 95% CI)	0.83 [0.33, 2.08]
2.3 Amitriptyline + perphenazine versus amitriptyline	1	41	Risk Ratio (M-H, Fixed, 95% CI)	1.73 [0.35, 8.41]

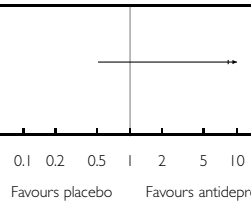
Analysis 1.1. Comparison 1 Antidepressant versus placebo, Outcome 1 Clinical response.

Review: Pharmacological treatment for psychotic depression

Comparison: 1 Antidepressant versus placebo

Outcome: 1 Clinical response

Study or subgroup	Antidepressant n/N	Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 Amitriptyline versus placebo				
Spiker 1988	4/14	0/13		8.40 [0.50, 142.27]

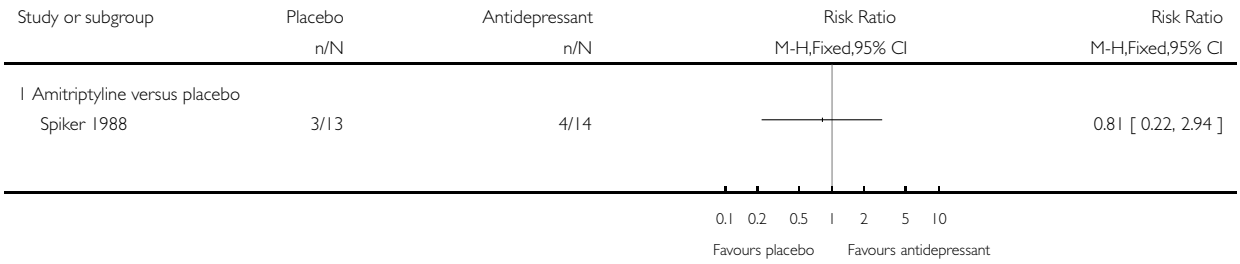


Analysis 1.2. Comparison 1 Antidepressant versus placebo, Outcome 2 Dropouts.

Review: Pharmacological treatment for psychotic depression

Comparison: 1 Antidepressant versus placebo

Outcome: 2 Dropouts

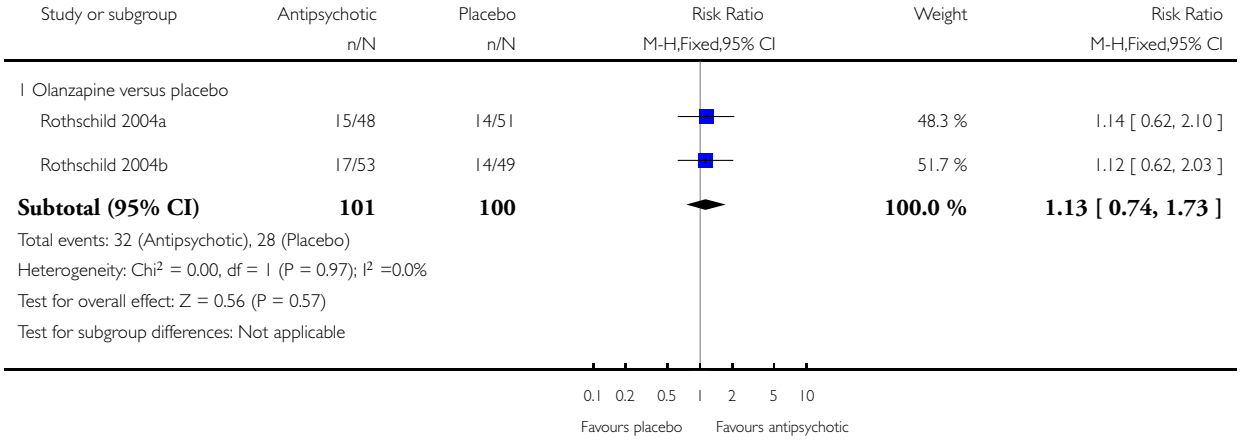


Analysis 2.1. Comparison 2 Antipsychotic versus placebo, Outcome 1 Clinical response.

Review: Pharmacological treatment for psychotic depression

Comparison: 2 Antipsychotic versus placebo

Outcome: 1 Clinical response

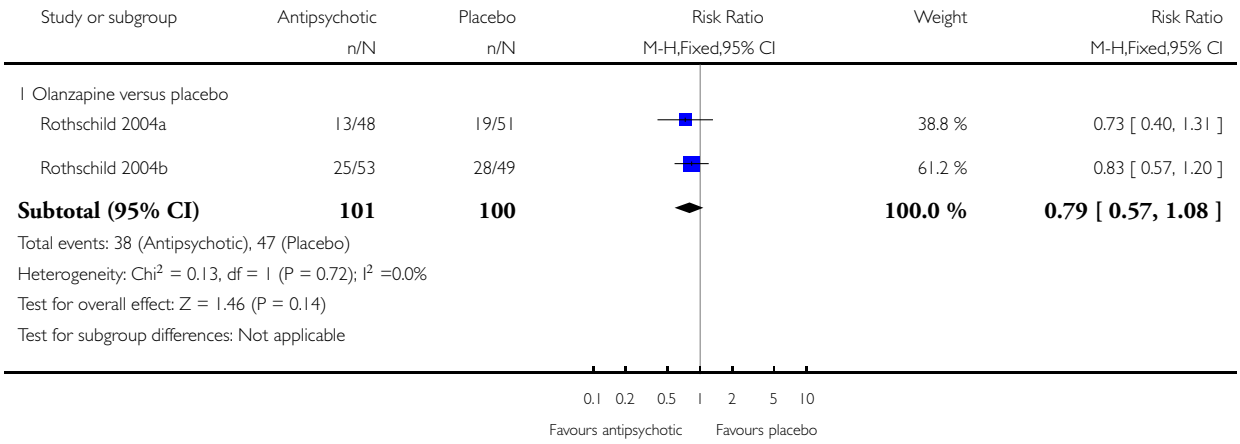


Analysis 2.2. Comparison 2 Antipsychotic versus placebo, Outcome 2 Dropouts.

Review: Pharmacological treatment for psychotic depression

Comparison: 2 Antipsychotic versus placebo

Outcome: 2 Dropouts

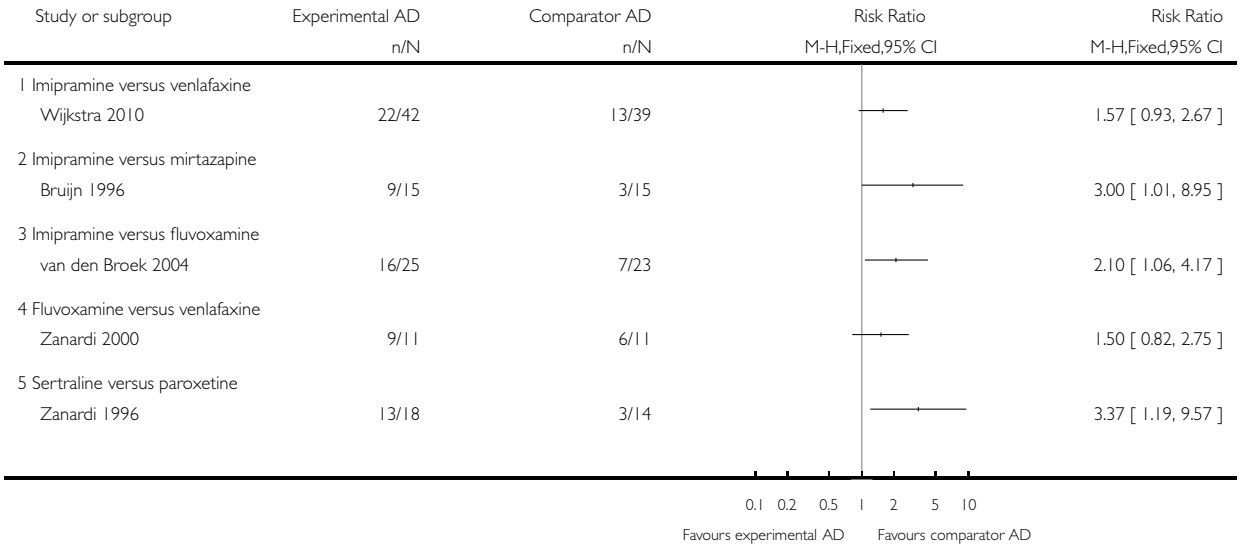


Analysis 3.1. Comparison 3 Antidepressant versus antidepressant, Outcome 1 Clinical response.

Review: Pharmacological treatment for psychotic depression

Comparison: 3 Antidepressant versus antidepressant

Outcome: 1 Clinical response



Analysis 3.2. Comparison 3 Antidepressant versus antidepressant, Outcome 2 Dropouts.

Review: Pharmacological treatment for psychotic depression

Comparison: 3 Antidepressant versus antidepressant

Outcome: 2 Dropouts

Study or subgroup	Favours experimental AD n/N	Comparator AD n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 Imipramine versus venlafaxine				
Wijkstra 2010	7/42	8/39		0.81 [0.33, 2.03]
2 Imipramine versus mirtazapine				
Bruijn 1996	4/15	8/15		0.50 [0.19, 1.31]
3 Imipramine versus fluvoxamine				
van den Broek 2004	4/25	2/25		2.00 [0.40, 9.95]
4 Fluvoxamine versus venlafaxine				
Zanardi 2000	0/11	2/11		0.20 [0.01, 3.74]
5 Sertraline versus paroxetine				
Zanardi 1996	0/18	5/14		0.07 [0.00, 1.20]

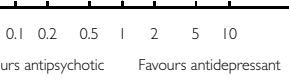
Analysis 4.1. Comparison 4 Antidepressant versus antipsychotic, Outcome 1 Clinical response.

Review: Pharmacological treatment for psychotic depression

Comparison: 4 Antidepressant versus antipsychotic

Outcome: 1 Clinical response

Study or subgroup	Antidepressant n/N	Antipsychotic n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 Amitriptyline versus perphenazine				
Spiker 1985	7/19	3/17		2.09 [0.64, 6.82]



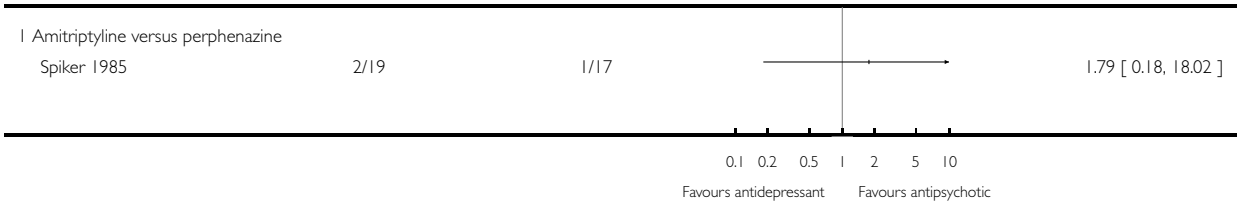
Analysis 4.2. Comparison 4 Antidepressant versus antipsychotic, Outcome 2 Dropouts.

Review: Pharmacological treatment for psychotic depression

Comparison: 4 Antidepressant versus antipsychotic

Outcome: 2 Dropouts

Study or subgroup	Antidepressant n/N	Antipsychotic n/N	Risk Ratio M-H,Fixed,95% CI	Risk Ratio M-H,Fixed,95% CI
1 Amitriptyline versus perphenazine				
Spiker 1985	2/19	1/17		1.79 [0.18, 18.02]

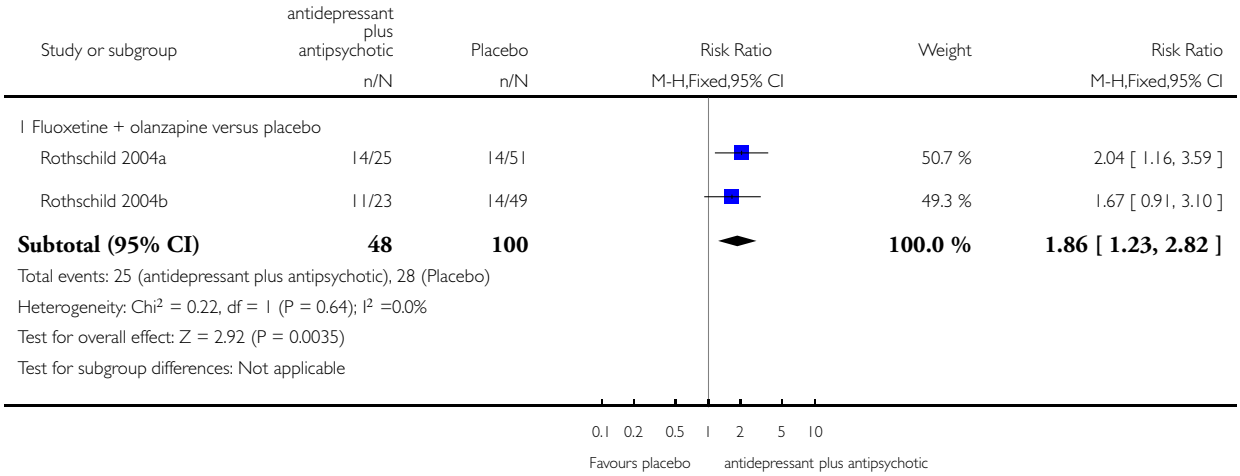


Analysis 5.1. Comparison 5 Antidepressant plus antipsychotic versus placebo, Outcome 1 Clinical response.

Review: Pharmacological treatment for psychotic depression

Comparison: 5 Antidepressant plus antipsychotic versus placebo

Outcome: 1 Clinical response

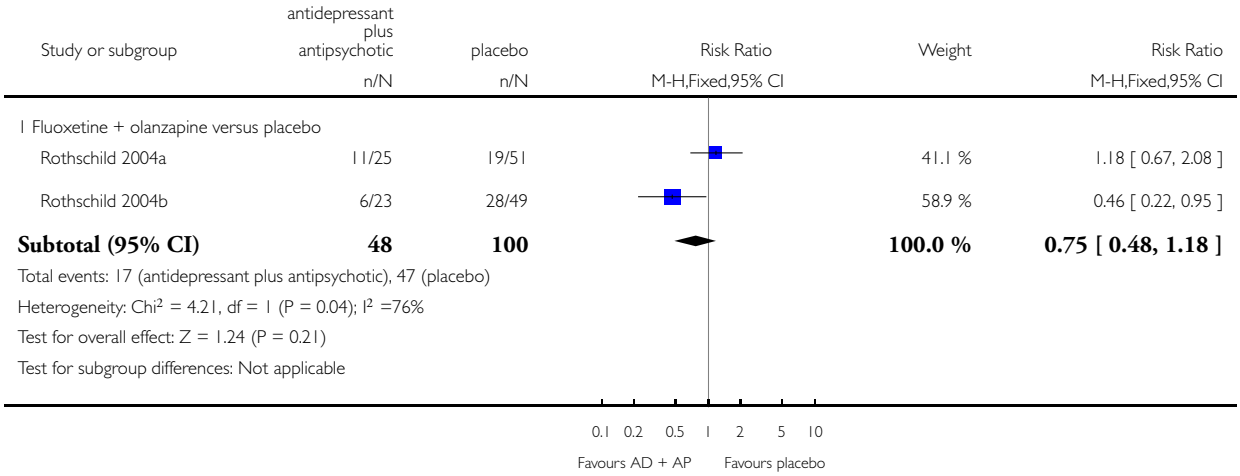


Analysis 5.2. Comparison 5 Antidepressant plus antipsychotic versus placebo, Outcome 2 Dropouts.

Review: Pharmacological treatment for psychotic depression

Comparison: 5 Antidepressant plus antipsychotic versus placebo

Outcome: 2 Dropouts

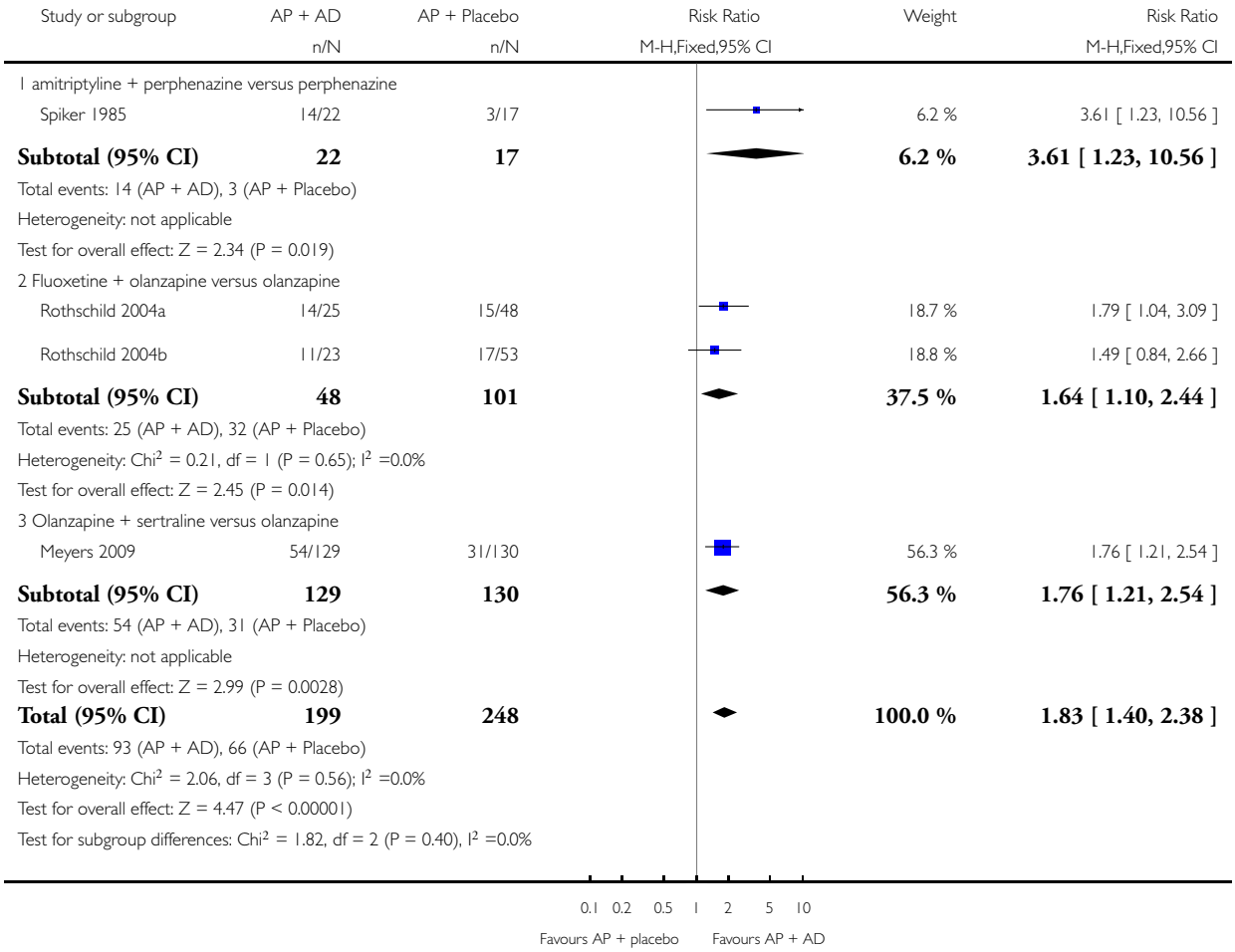


Analysis 6.1. Comparison 6 Antidepressant plus antipsychotic versus placebo plus antipsychotic, Outcome 1 Clinical response.

Review: Pharmacological treatment for psychotic depression

Comparison: 6 Antidepressant plus antipsychotic versus placebo plus antipsychotic

Outcome: 1 Clinical response

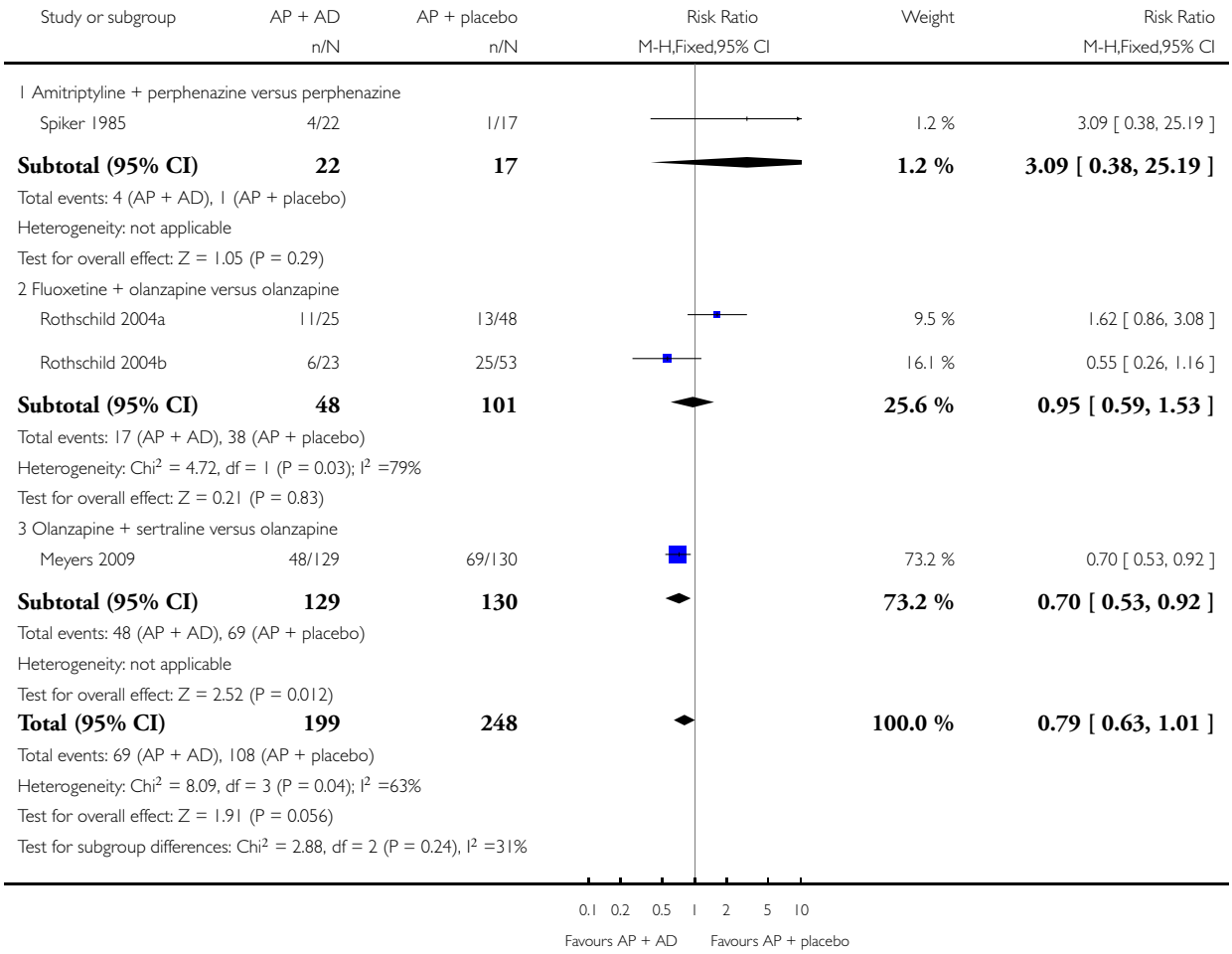


Analysis 6.2. Comparison 6 Antidepressant plus antipsychotic versus placebo plus antipsychotic, Outcome 2 Dropouts.

Review: Pharmacological treatment for psychotic depression

Comparison: 6 Antidepressant plus antipsychotic versus placebo plus antipsychotic

Outcome: 2 Dropouts

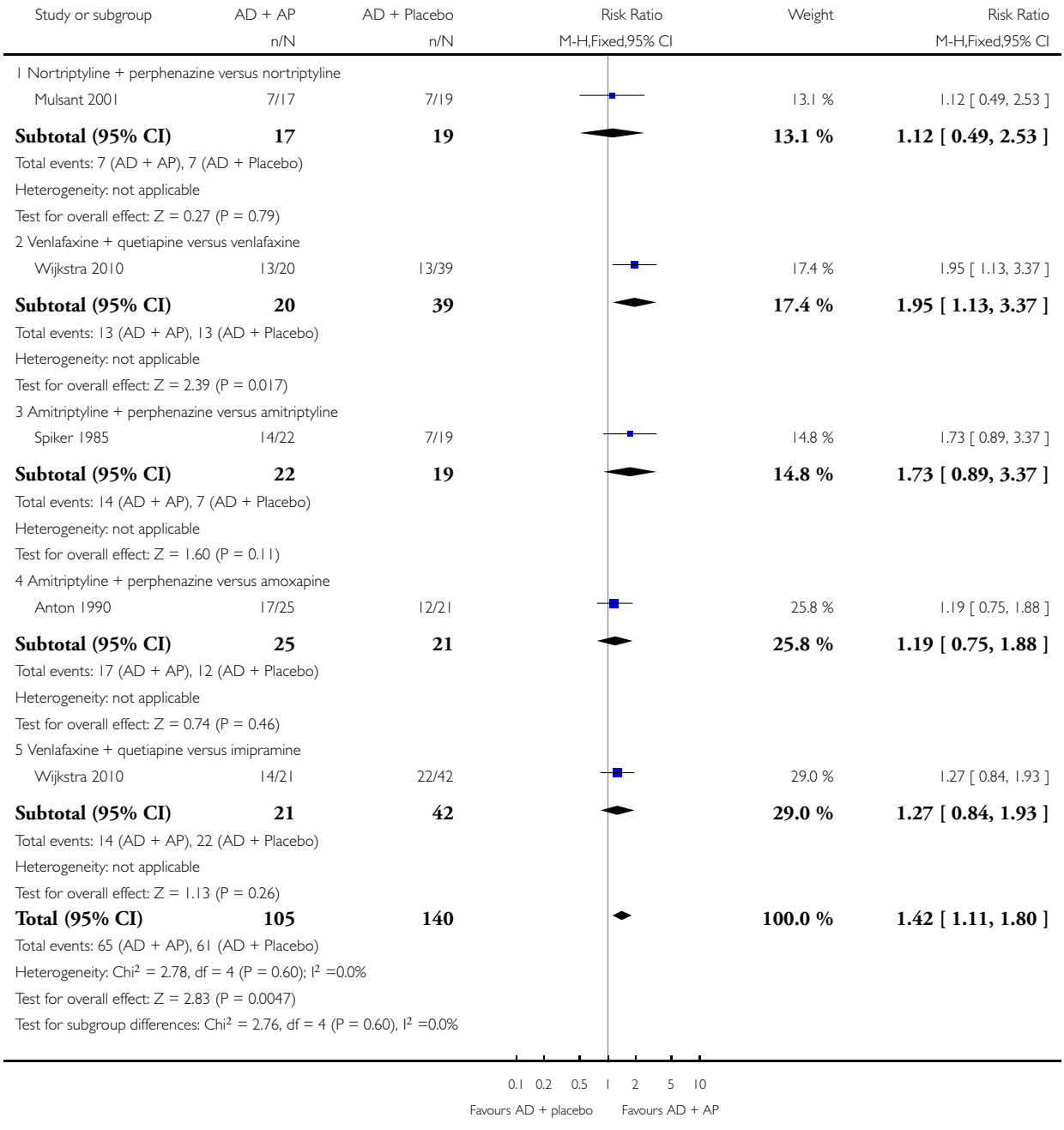


Analysis 7.1. Comparison 7 Antidepressant plus antipsychotic versus placebo plus antidepressant, Outcome 1 Clinical response.

Review: Pharmacological treatment for psychotic depression

Comparison: 7 Antidepressant plus antipsychotic versus placebo plus antidepressant

Outcome: 1 Clinical response

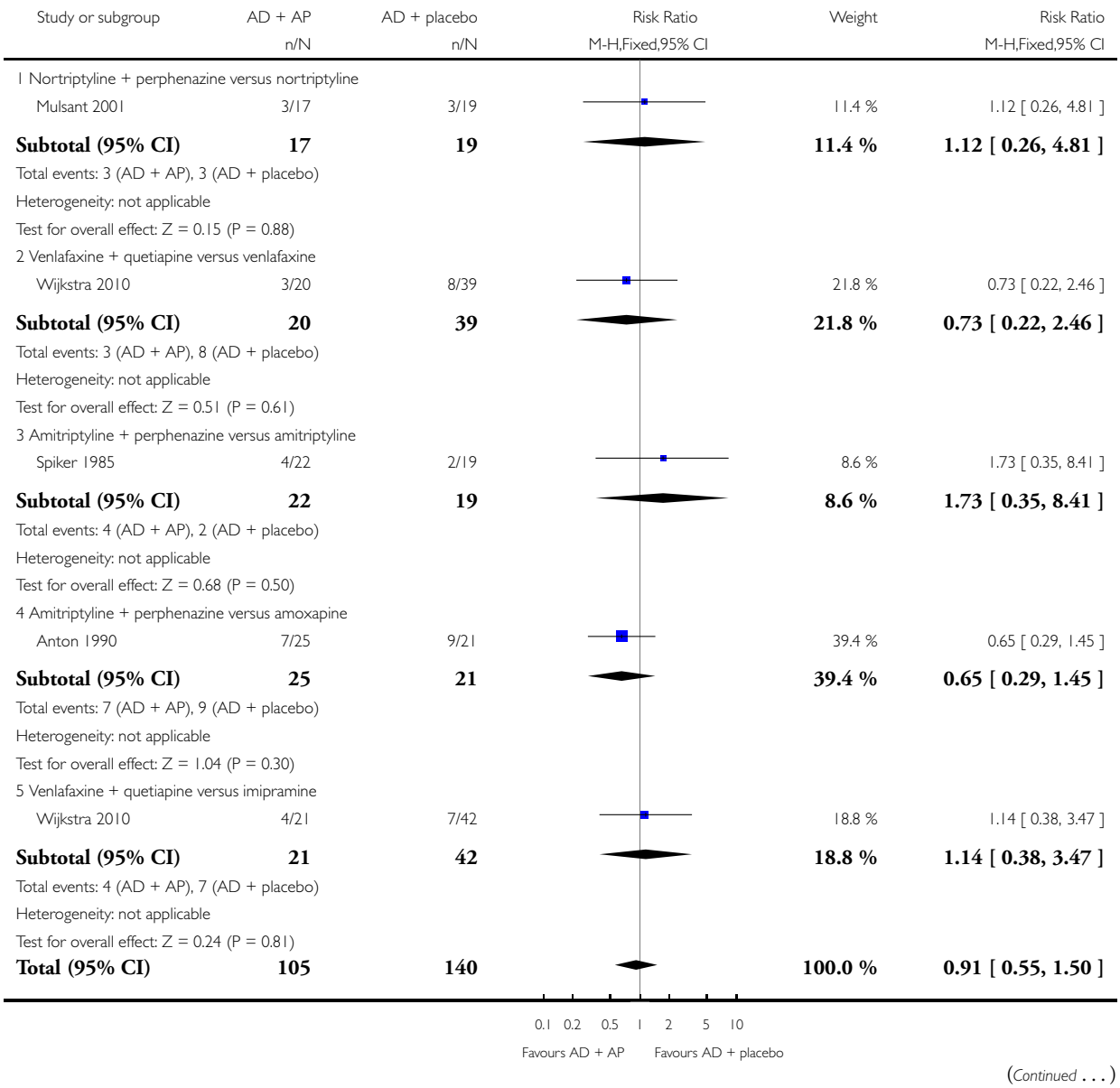


Analysis 7.2. Comparison 7 Antidepressant plus antipsychotic versus placebo plus antidepressant, Outcome 2 Dropouts.

Review: Pharmacological treatment for psychotic depression

Comparison: 7 Antidepressant plus antipsychotic versus placebo plus antidepressant

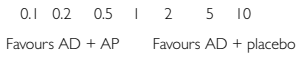
Outcome: 2 Dropouts



(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	AD + AP n/N	AD + placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Total events: 21 (AD + AP), 29 (AD + placebo)					
Heterogeneity: Chi^2 = 1.65, df = 4 (P = 0.80); I^2 =0.0%					
Test for overall effect: Z = 0.38 (P = 0.71)					
Test for subgroup differences: Chi^2 = 1.64, df = 4 (P = 0.80), I^2 =0.0%					



Analysis 8.1. Comparison 8 Antidepressant plus antipsychotic versus placebo plus the same antidepressant, Outcome 1 Clinical response.

Review: Pharmacological treatment for psychotic depression

Comparison: 8 Antidepressant plus antipsychotic versus placebo plus the same antidepressant

Outcome: 1 Clinical response

Study or subgroup	AD + AP n/N	AD + Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
1 Nortriptyline + perphenazine versus nortriptyline					
Mulsant 2001	7/17	7/19		24.1 %	1.12 [0.49, 2.53]
Subtotal (95% CI)	**17**	**19**		**24.1 %**	**1.12 [0.49, 2.53]**
Total events: 7 (AD + AP), 7 (AD + Placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.27 (P = 0.79)					
2 Venlafaxine + quetiapine versus venlafaxine					
Wijkstra 2010	27/41	13/39		48.5 %	1.98 [1.20, 3.24]
Subtotal (95% CI)	**41**	**39**		**48.5 %**	**1.98 [1.20, 3.24]**
Total events: 27 (AD + AP), 13 (AD + Placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 2.69 (P = 0.0071)					
3 Amitriptyline + perphenazine versus amitriptyline					
Spiker 1985	14/22	7/19		27.4 %	1.73 [0.89, 3.37]
Subtotal (95% CI)	**22**	**19**		**27.4 %**	**1.73 [0.89, 3.37]**
Total events: 14 (AD + AP), 7 (AD + Placebo)					
Heterogeneity: not applicable					

(*Continued* . . .)

(. . . *Continued*)

Study or subgroup	AD + AP n/N	AD + Placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
Test for overall effect: Z = 1.60 (P = 0.11)					
Total (95% CI)	**80**	**77**		**100.0 %**	**1.70 [1.19, 2.43]**
Total events: 48 (AD + AP), 27 (AD + Placebo)					
Heterogeneity: Chi^2 = 1.36, df = 2 (P = 0.51); I^2 =0.0%					
Test for overall effect: Z = 2.93 (P = 0.0034)					
Test for subgroup differences: Chi^2 = 1.36, df = 2 (P = 0.51), I^2 =0.0%					

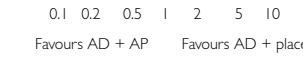
Analysis 8.2. Comparison 8 Antidepressant plus antipsychotic versus placebo plus the same antidepressant, Outcome 2 Dropouts.

Review: Pharmacological treatment for psychotic depression

Comparison: 8 Antidepressant plus antipsychotic versus placebo plus the same antidepressant

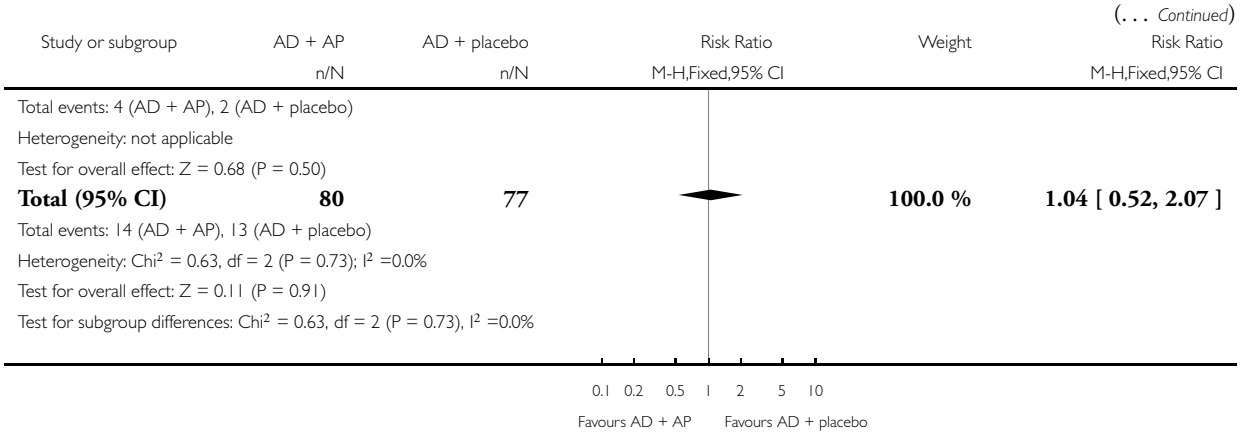
Outcome: 2 Dropouts

Study or subgroup	AD + AP n/N	AD + placebo n/N	Risk Ratio M-H,Fixed,95% CI	Weight	Risk Ratio M-H,Fixed,95% CI
1 Nortriptyline + perphenazine versus nortriptyline					
Mulsant 2001	3/17	3/19		21.5 %	1.12 [0.26, 4.81]
Subtotal (95% CI)	**17**	**19**		**21.5 %**	**1.12 [0.26, 4.81]**
Total events: 3 (AD + AP), 3 (AD + placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.15 (P = 0.88)					
2 Venlafaxine + quetiapine versus venlafaxine					
Wijkstra 2010	7/41	8/39		62.2 %	0.83 [0.33, 2.08]
Subtotal (95% CI)	**41**	**39**		**62.2 %**	**0.83 [0.33, 2.08]**
Total events: 7 (AD + AP), 8 (AD + placebo)					
Heterogeneity: not applicable					
Test for overall effect: Z = 0.39 (P = 0.69)					
3 Amitriptyline + perphenazine versus amitriptyline					
Spiker 1985	4/22	2/19		16.3 %	1.73 [0.35, 8.41]
Subtotal (95% CI)	**22**	**19**		**16.3 %**	**1.73 [0.35, 8.41]**



(*Continued* . . .)

(. . . *Continued*)



APPENDICES

Appendix 1. CENTRAL update search 2010

The Cochrane Register of Controlled Trails (CENTRAL) was searched (Issue 4, 2010) using the following terms:
#1 MeSH descriptor DEPRESSION, this term only
#2 MeSH descriptor DEPRESSIVE DISORDER, this term only
#3 MeSH descriptor DEPRESSIVE DISORDER MAJOR, this term only
#4 (depression* or depressive*):ti,ab,kw
#5 (#1 or #2 or #3 or #4)
#6 MeSH descriptor DELUSIONS, this term only
#7 delusion*:ti,ab,kw
#8 MeSH descriptor PSYCHOTIC DISORDERS, this term only
#9 MeSH AFFECTIVE DISORDERS, PSYCHOTIC, this term only
#10 (psychotic* or psychosis or psychoses) :ti,ab,kw
#11 (#6 or #7 or #8 or #9 or #10)
#12 (#5 and #11), from 2005 to 2010
#13 SR-DEPRESSN or HS-DEPRESSN
#14 (#12 NOT #13)

Appendix 2. Previous search strategies to 2005

We searched the Cochrane Central Register of Controlled Trials (CENTRAL) with the terms depressive disorder and drug treatment. In addition we searched MEDLINE (1966 until April 2004) and EMBASE (1980 until April 2004) using the following terms: ("depressive disorder/drug therapy"[MESH] AND (("delusions"[MESH Terms] OR delusions[Text Word]) OR (("psychotic disorders"[MESH Terms] OR psychotic[Text Word]) AND features[All Fields])))) combined with a sensitive search strategy for RCTs.

FEEDBACK

Feedback submitted, 3 February 2015

Summary

We found a possible error in the review 'Pharmacological treatment for psychotic depression' by Wijkstra J, Lijmer J, Burger H, Geddes J, Nolen WA., which was published in issue 11 of year 2013.

When we read through the review, we found that they included 2 comparisons from a single article to calculate clinical outcome in their Analysis 7. The referenced article was Wijkstra 2010, in which 122 patients were randomized into 3 treatment groups; imipramine (n= 42), venlafaxine (n=39) or venlafaxine + quetiapine (n=41). In their Analysis 7, they compared imipramine or venlafaxine group against venlafaxine + quetiapine group independently in each subgroup. Then, when they conducted the analysis for the Total, venlafaxine + quetiapine group (n=41) was included twice in the "antidepressant plus antipsychotic" group.

Double counting the same subjects would spuriously increase precision in the meta-analytic estimates. Authors should use a proper method to avoid double-counting the same subjects.

Reply

We would like to thank Dr Matsuo and his colleagues for pointing out this mistake in the original analysis. We looked at this and we agreed that the best approach is probably to split the comparator (as described in the Cochrane Handbook, chapter 16.5.4). We amended the analyses in the revised review accordingly and, given the numbers involved, it makes no material difference to the point estimates or precision. The revised estimate for clinical response (see Analysis 7.1) was: RR 1.42; 95% CI 1.11 to 1.81 (while the original pooled RR was 1.44 with a 95% CI 1.15 to 1.80). The revised estimate for dropouts (see Analysis 7.2) was: RR 0.91, 95% CI 0.55 to 1.50 (the original pooled RR was 0.91, with a 95% CI 0.58 to 1.44).

We thank the EBMH Study Group for their interest and close reading of our review.

Contributors

Feedback submitted by: Masahiro Matsuo, Aran Tajika, Toshi A. Furukawa, Kyoto EBMH Study Group
Response submitted by: Andrea Cipriani and John Geddes

WHAT'S NEW

Last assessed as up-to-date: 12 April 2013.

Date	Event	Description
9 July 2015	Feedback has been incorporated	Mistake in original analysis was corrected. This made no material difference to the results
10 June 2013	New search has been performed	Searches and methodology updated
10 June 2013	New citation required and conclusions have changed	Update of previous review. Two new studies included; conclusions slightly revised

HISTORY

Protocol first published: Issue 1, 2003

Review first published: Issue 4, 2005

Date	Event	Description
2 September 2010	Amended	Methods updated to reflect current Handbook
3 November 2008	Amended	Converted to new review format.
10 August 2005	New citation required and conclusions have changed	Substantive amendment

CONTRIBUTIONS OF AUTHORS

J Wijkstra: development of protocol, coordination and writing of the review, data collection, analysis, primary author report.

J Lijmer: development of protocol, data collection, analysis, co-author report.

H Burger: statistical advice.

J Geddes: co-author report.

WA Nolen: development of protocol, data collection, analysis, overall supervision, co-author report.

A Cipriani: statistical advice and analysis.

DECLARATIONS OF INTEREST

JW and WN conducted a multi-centre trial in participants with psychotic depression that compared treatment with imipramine, venlafaxine and venlafaxine plus quetiapine. Wyeth and AstraZeneca financially supported this trial. Data from this trial are included in this review. To prevent bias, the data extracted from our own study explicitly have been checked by the Cochrane organisation.

JG has received research funding and support from sanofi-aventis and GlaxoSmithKline and is currently in discussion with several other companies that manufacture SSRIs about collaboration on planned independent trials and systematic reviews.

This review has been undertaken without external support.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

All studies were evaluated according to the new method used for assessing risk of bias. The background section has been updated.

INDEX TERMS

Medical Subject Headings (MeSH)

Antidepressive Agents [*therapeutic use]; Antipsychotic Agents [*therapeutic use]; Depressive Disorder, Major [*drug therapy; etiology]; Drug Therapy, Combination [methods]; Psychotic Disorders [complications; *drug therapy]; Randomized Controlled Trials as Topic

MeSH check words

Humans